

## College of Information Science and Technology



Drexel E-Repository and Archive (iDEA)  
<http://idea.library.drexel.edu/>

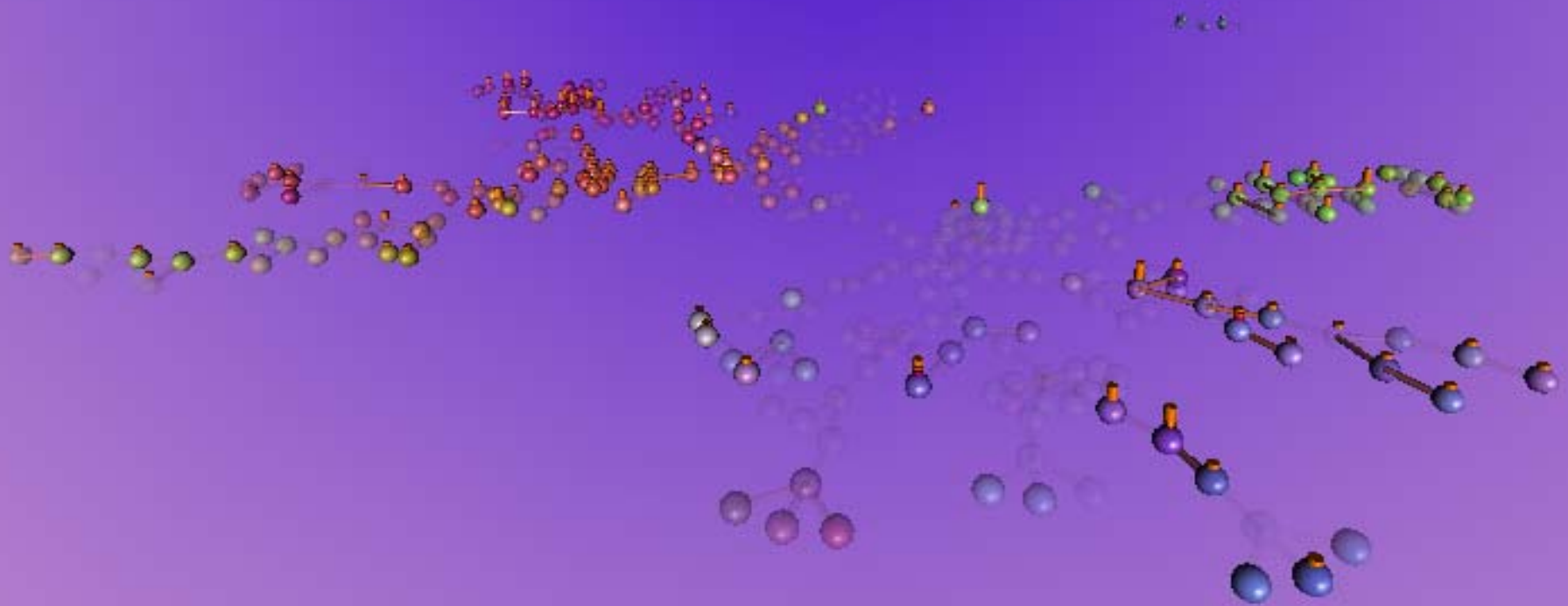
Drexel University Libraries  
[www.library.drexel.edu](http://www.library.drexel.edu)

The following item is made available as a courtesy to scholars by the author(s) and Drexel University Library and may contain materials and content, including computer code and tags, artwork, text, graphics, images, and illustrations (Material) which may be protected by copyright law. Unless otherwise noted, the Material is made available for non profit and educational purposes, such as research, teaching and private study. For these limited purposes, you may reproduce (print, download or make copies) the Material without prior permission. All copies must include any copyright notice originally included with the Material. **You must seek permission from the authors or copyright owners for all uses that are not allowed by fair use and other provisions of the U.S. Copyright Law.** The responsibility for making an independent legal assessment and securing any necessary permission rests with persons desiring to reproduce or use the Material.

Please direct questions to [archives@drexel.edu](mailto:archives@drexel.edu)

# Visual Analysis of Concept Change and Information Diffusion

Chaomei Chen, Drexel University



# Understanding the Structure, the Growth, and Spread of Information and Knowledge

- How often do we need to answer **these** questions?
  - (as a scientist, a scholar, an evaluator, a policy maker, a doctoral student, a librarian, etc.)
- What tools are available for such needs?
- How effective are they?

1. What is the hottest topic at time  $T_0$ ?
2. What have been the hot topics between time  $T_a$  and  $T_b$ ? (A Timeline - Temporal)
3. What are the major turning points between time  $T_a$  and  $T_b$ ? (Turning points - Structural)
4. How did knowledge associated with these turning points spread? (Diffusion - Spatial)
5. How can we differentiate various perspectives and views based on a large volume of inputs? (Sense-Making - Semantics)

# A quick CiteSpace demo

Topic: **Terrorism**

Time span: **1996-2003**

**Q1: Current hot topic?**

1. What is the most studied topic in the area of **TERRORISM RESEARCH**?

**Q2: Previous hot topic?**

2. What topic was predominant before this one?

**Q3: Turning point?**

3. What **EVENT/WORK** caused the shift of focus?

**Q4: Transition path?**

4. What is the intellectual path that underlined such a transition?



## Projects

New Edit Delete Terrorism 1990-2006

Project Home: C:\CiteSpaceProjects\Terroris--1990-2006

Data Directory: C:\DATA\Terroris- 1990-2006.3.21

GO! Stop Reset JVM Memory (K) 509927 Used (%) 51

## Space Status

1-year slices	c   cc   ccv	space	nodes	links
Pruning configuration:				
1996-1996	5   3   0.1	1851	1	0
1997-1997	5   3   0.12	2546	5	10
1998-1998	5   3   0.15	2010	1	0
1999-1999	5   3   0.17	2117	15	39
2000-2000	5   3   0.2	3059	7	10
2001-2001	5   3   0.2	4930	38	189
2002-2002	5   3   0.2	9188	30	65
2003-2003	5   3   0.2	18171	80	317

## Process Reports

download4401-4700.txt  
 download4701-5000.txt  
 download5001-5300.txt  
 download5301-5500.txt  
 download5501-5698.txt  
 download601-900.txt  
 download901-1200.txt

Run time: 22256 milliseconds.

Merged network: Nodes=122, Links=551

Exclusion List: 0

Network modeling ends at Wed Oct 25 20:10:04 EDT 2006.

## Time Slicing

From 1996 To 2003 # Years per Slice 1

## Modeling

## Nodes

- ☐ Citing Authors
 ☐ Citing Terms
 ☐ Cited Authors
 ☒ Cited Documents
 ☐ Cited Journals

## Term Sources

☒ Title
 ☒ Abstract
 ☒ Descriptors
 ☒ Identifiers

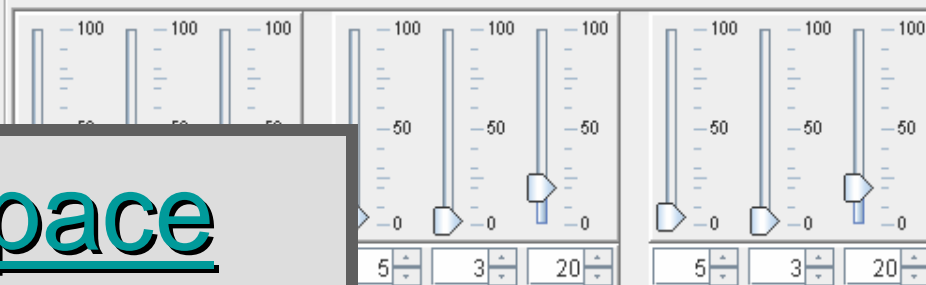
## Term Selection

☐ None
 ☐ Noun Phrases
 ☒ Burst Terms
 Find Burst Terms

## Links

Strength Cosine Scope Within Slices

## Thresholding (c, cc, ccv)



- ☐ Pruning sliced networks  
☐ Minimum Spanning Tree  
☐ Pruning the merged network

## Visualization

- ☒ Cluster View (Default)
 ☐ Time-zone View
 ☐ Show Networks by Time Slices
 ☒ Show Merged Network

# CiteSpace WebStart

# Finding Tipping Points

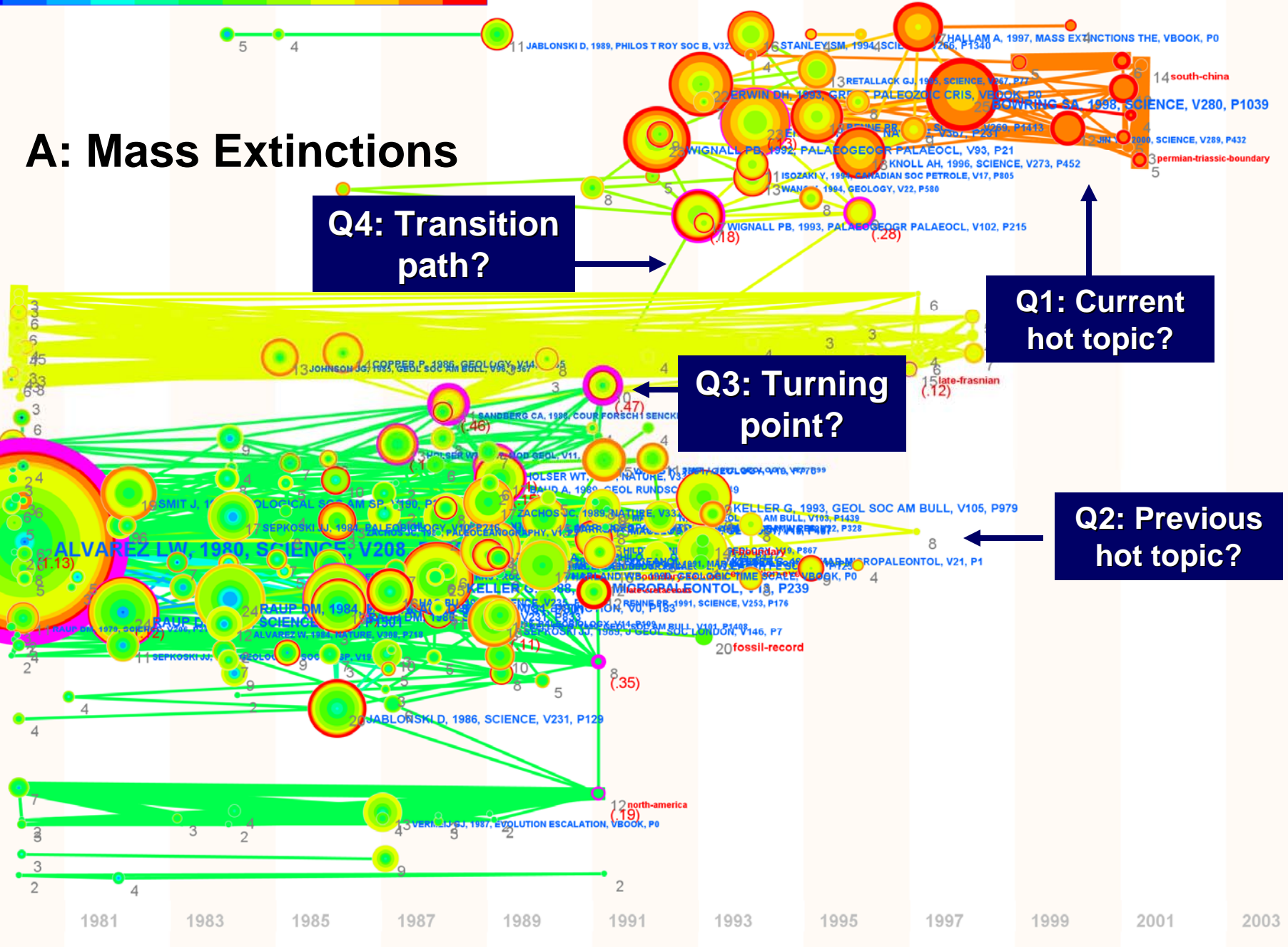
# A: Mass Extinctions

Q4: Transition path?

Q1: Current hot topic?

Q3: Turning point?

Q2: Previous hot topic?

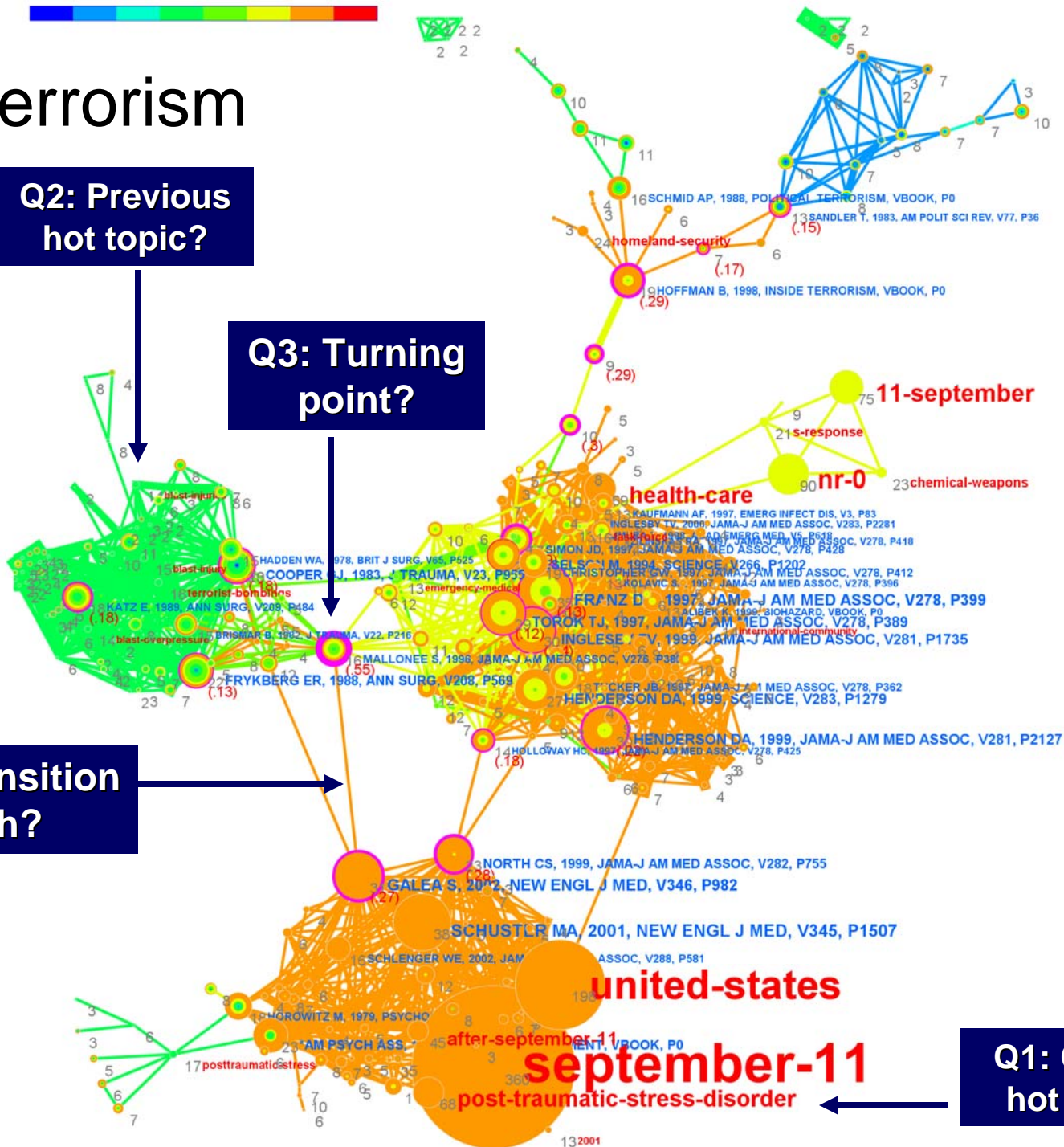


## Q2: Previous hot topic?

### Q3: Turning point?

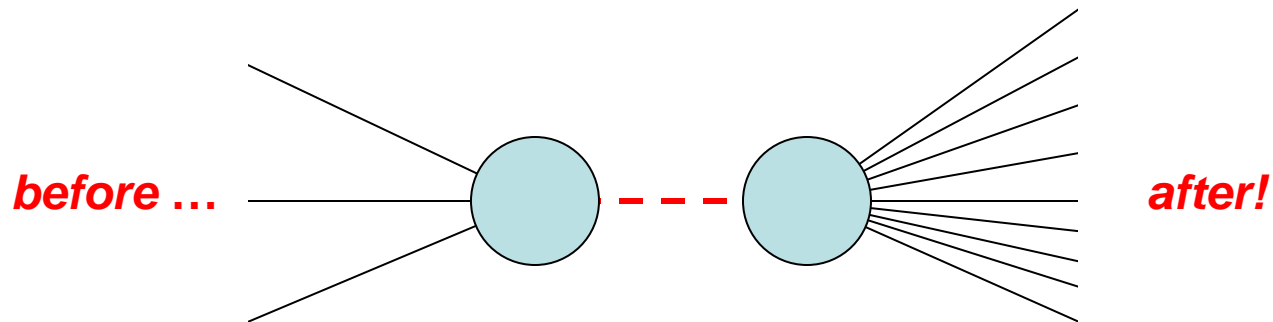
#### Q4: Transition path?

**Q1: Current hot topic?**



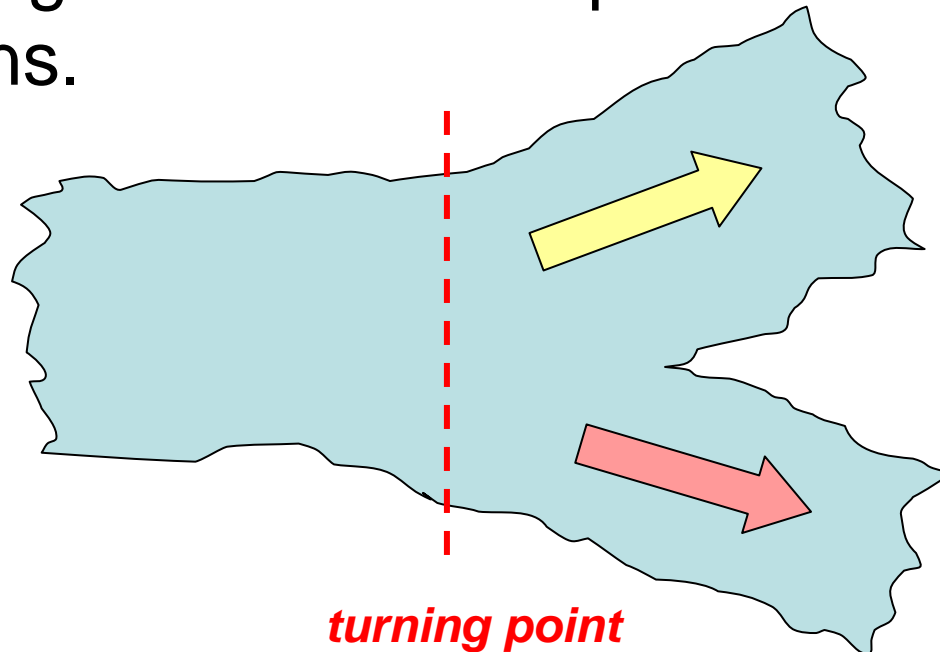
# Structural and Temporal Analysis of Scientific Literature

- Emerging themes
- Intellectual turning points

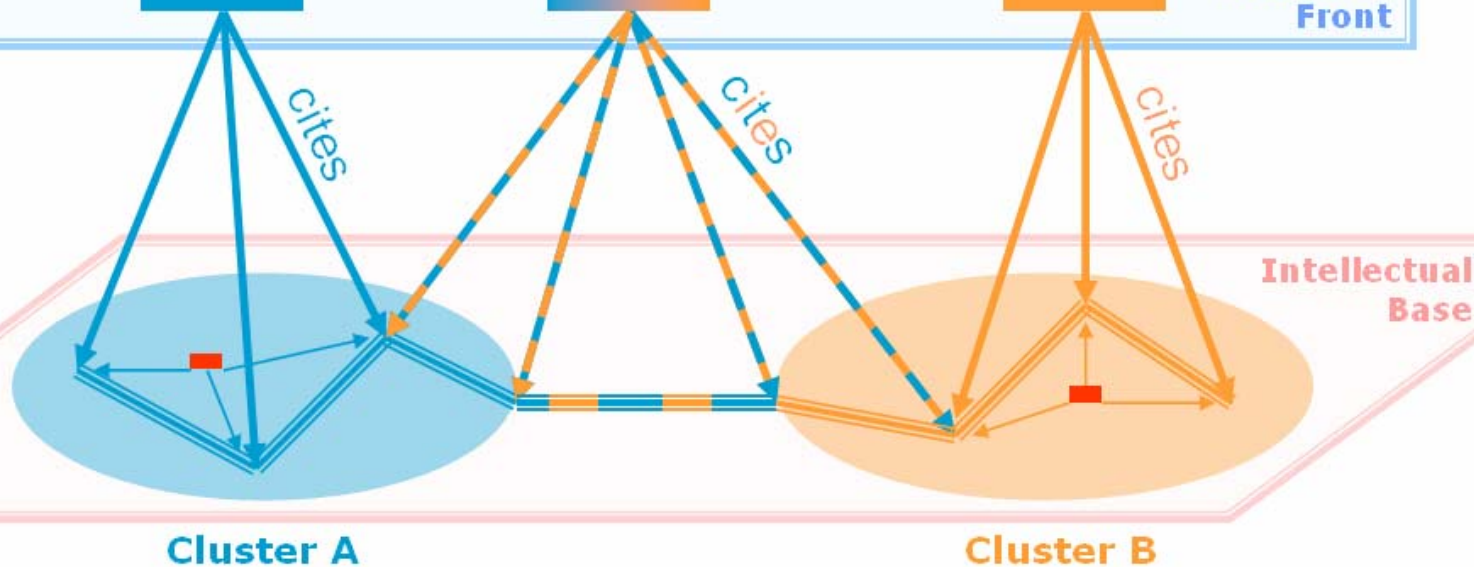


# Paradigm Shift

- Normative
  - Citations reflect intellectual values.
- Recentness
  - Citations register new concepts and new associations.



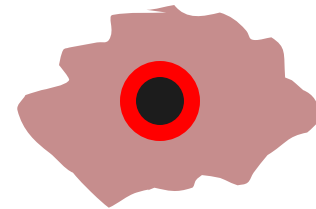
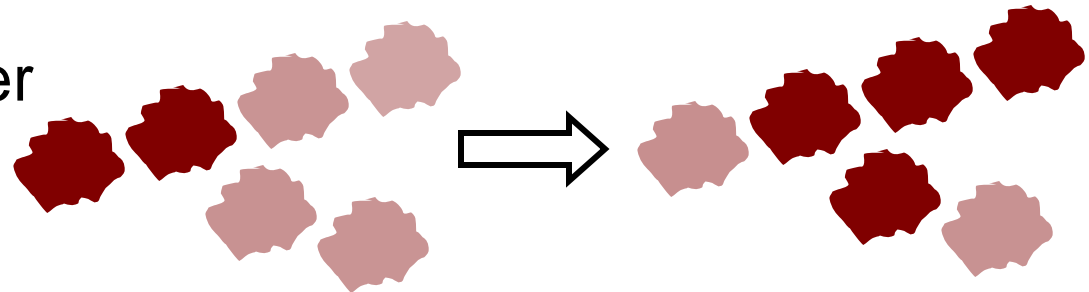
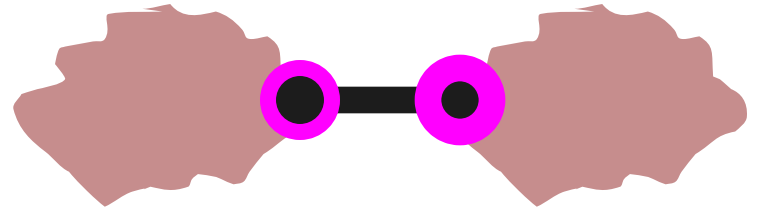
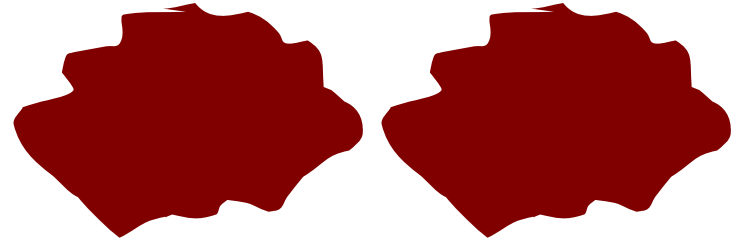




**Cross-cluster co-citation**  
**Within cluster co-citation**  
**Citing phrase**  
**Phrase-to-article citation**

# Expected Patterns

- Thematic grouping
- Intellectual turning points
- Thematic change over time
- Abrupt changes associated with triggers

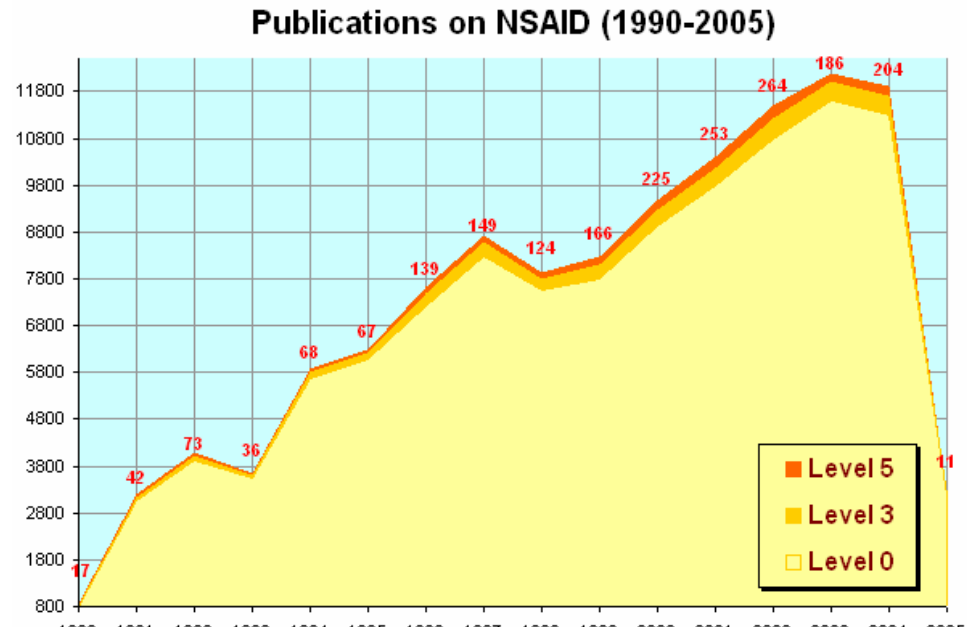


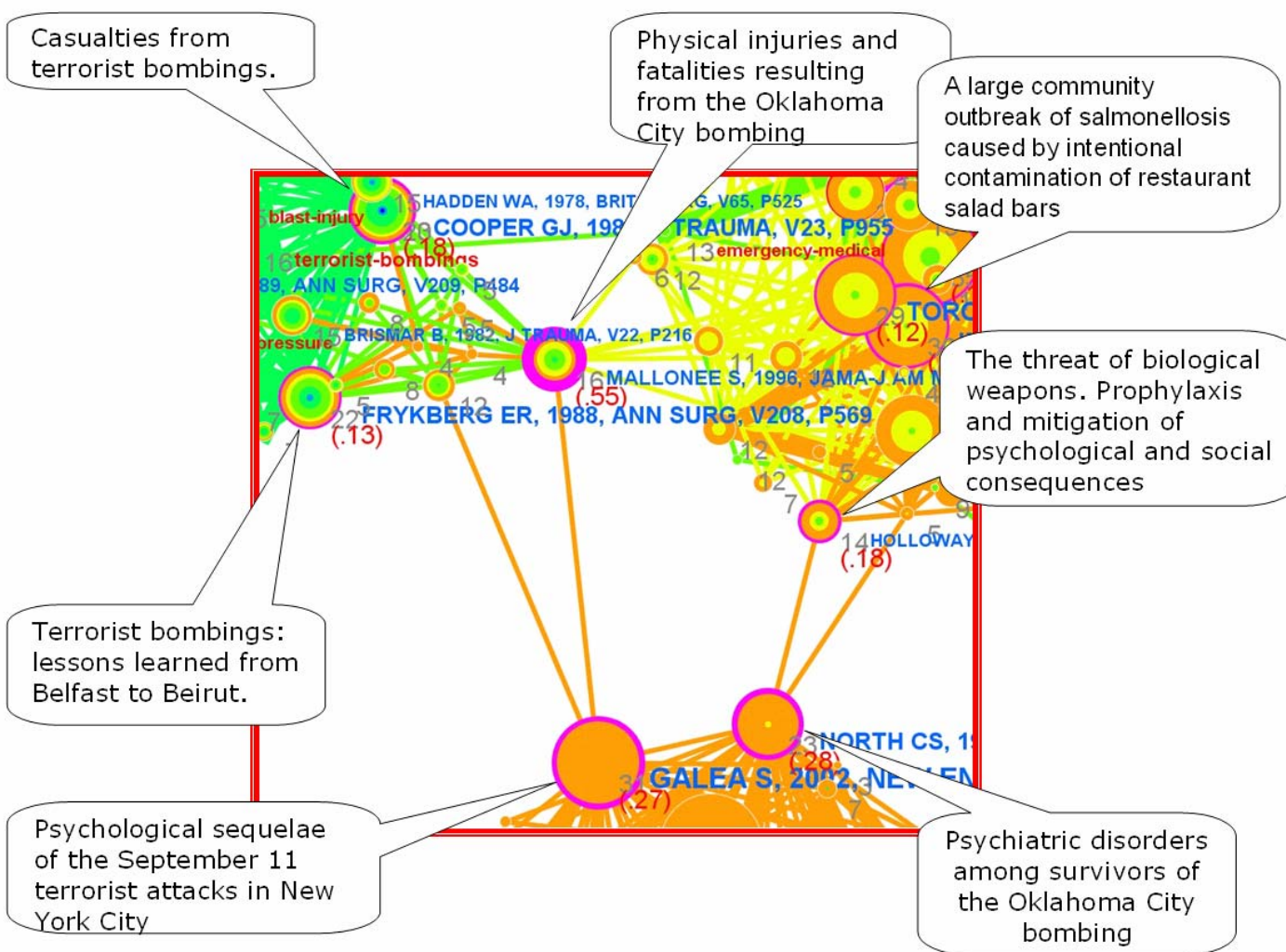


# Why Scientists Cite?

- Normative View

- Citations are made because of the intellectual values of cited works.
- They should not be affected by social and cultural characteristics such as race, gender, or academic rank.







11 Stress Disorders, Post-Traumatic/\*epidemiology

9 Terrorism/\*psychology

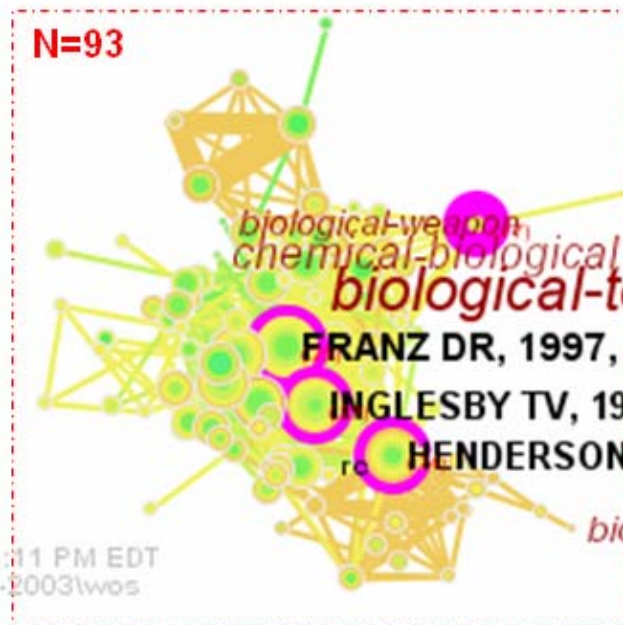
8 Disasters

27 Biological Warfare

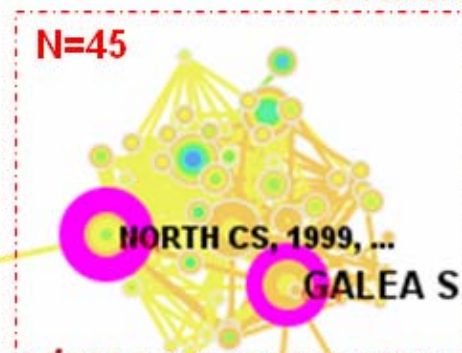
17 Violence

14 Bioterrorism

N=93



N=45



N=31



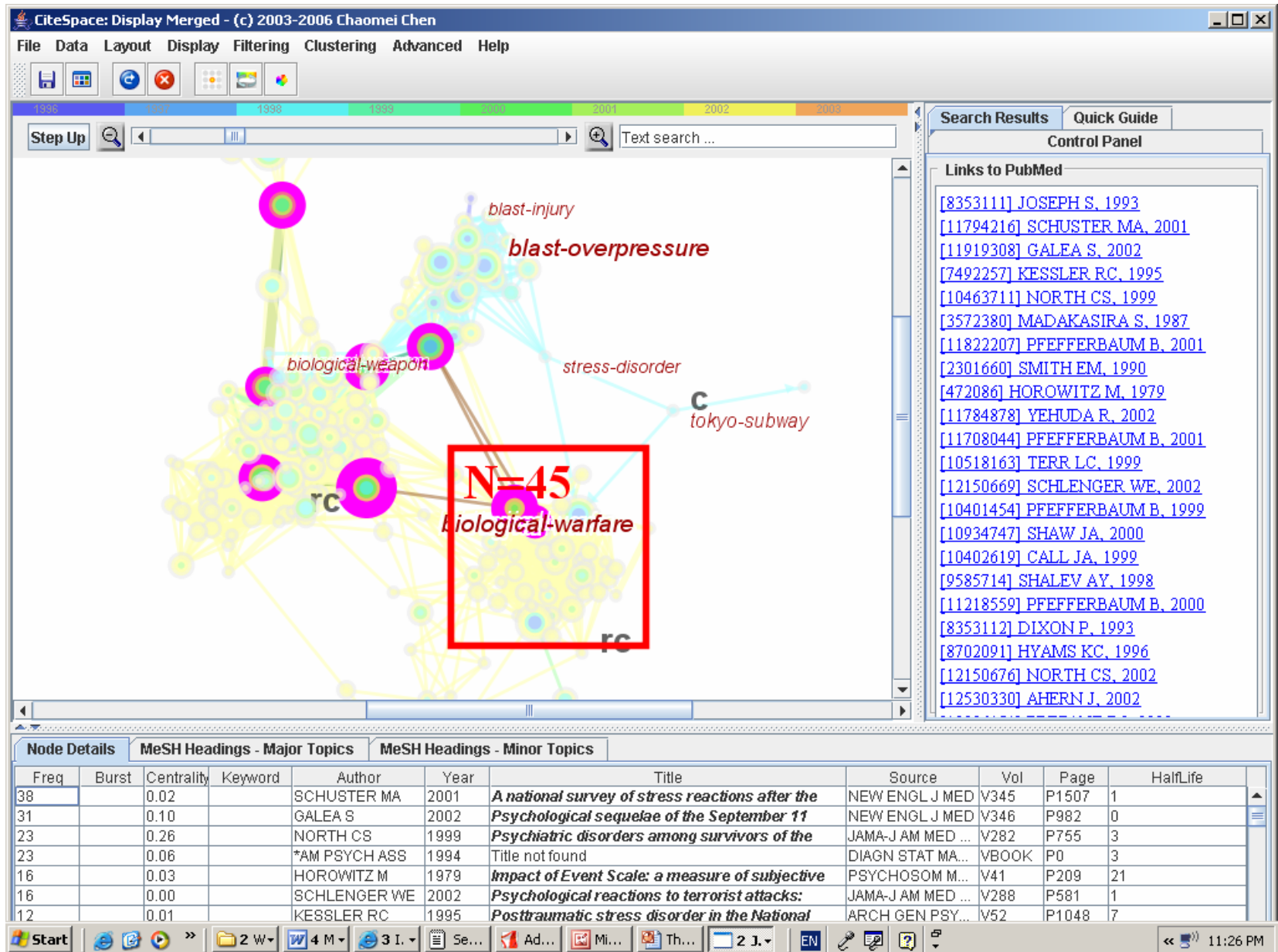
11 Explosions

7 Violence

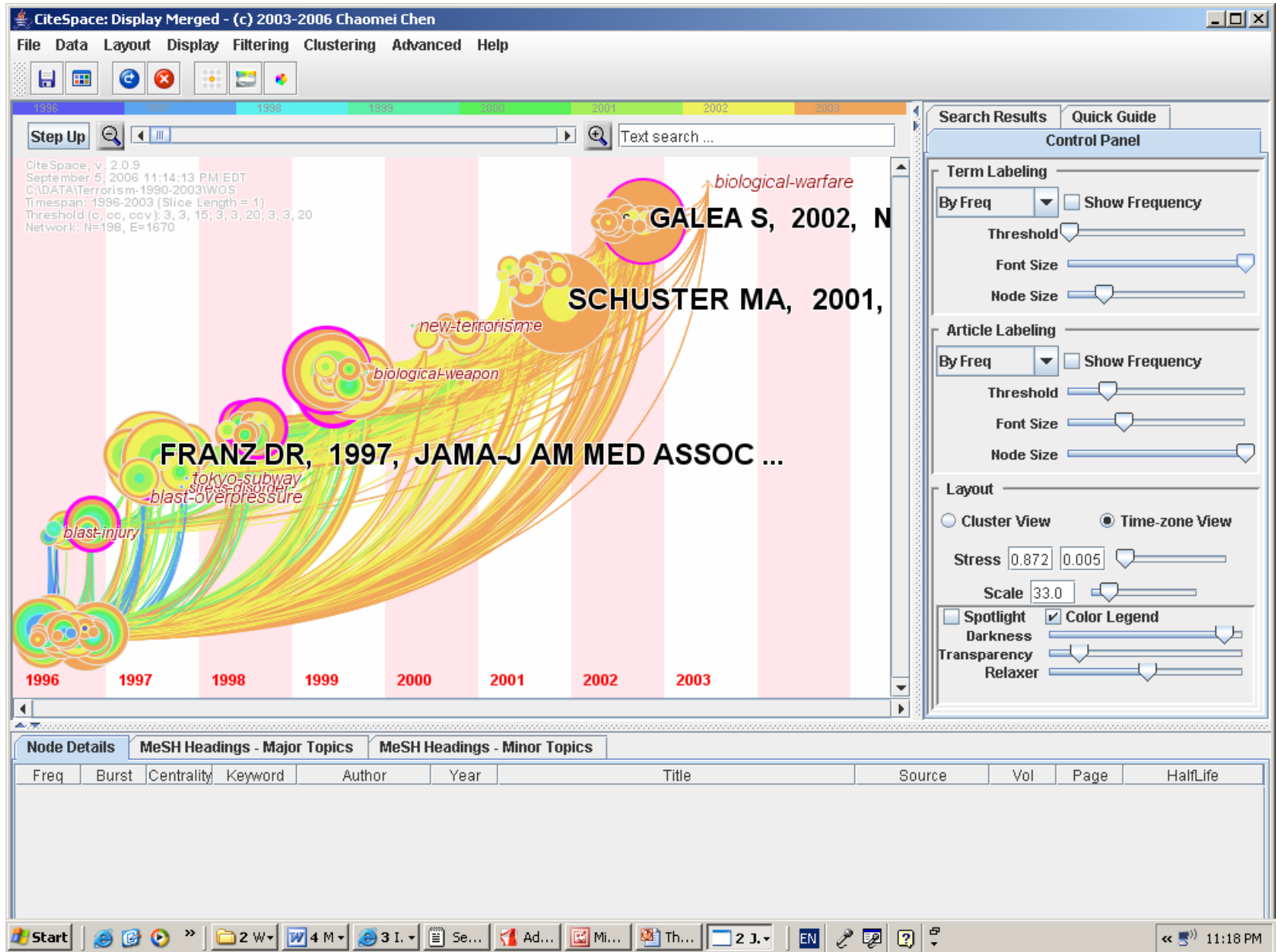
4 Blast Injuries/\*mortality

CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

# Cluster View



# Timezone View





# Users from 36 Countries Download 2+

(9/13/2004-8/23/2006)

#	Country	Unique IP	Downloads	#	Country	Unique IP	Downloads
1	USA	132	470	19	Austria	3	10
2	Spain	121	509	20	Mexico	3	10
3	Germany	35	101	21	Ireland	2	5
4	China	25	132	22	Israel	1	12
5	United Kingdom	16	64	23	Argentina	1	11
6	Taiwan	15	47	24	Kenya	1	7
7	Korea	12	40	25	Croatia	1	3
8	Canada	11	26	26	Finland	1	3
9	Denmark	10	62	27	India	1	3
10	The Netherlands	10	34	28	Puerto Rico	1	3
11	France	7	22	29	Slovenia	1	3
12	Japan	7	18	30	Belgium	1	2
13	Chile	6	21	31	Bulgaria	1	2
14	Russia	5	68	32	Egypt	1	2
15	Italy	4	23	33	Hong Kong	1	2
16	Brazil	4	16	34	Norway	1	2
17	Australia	4	14	35	Switzerland	1	2
18	Sweden	4	9	36	Thailand	1	2
					<b>Subtotal (downloads &gt;1)</b>	<b>451</b>	<b>1,760</b>
					<b>Grand Total</b>	<b>1,247</b>	<b>2,619</b>

# Users from 14 Countries WebStart It 30+

(9/13/2004-8/23/2006)

#	Country	Unique IP (WebStarts > 30)	Launches (#)
1	Spain	13	680
2	China	4	590
3	USA	7	351
4	Drexel	3	329
5	Denmark	2	142
6	United Kingdom	2	124
7	Canada	3	114
8	Italy	2	78
9	Chile	1	60
10	Taiwan	1	41
11	Argentina	1	40
12	Mexico	1	39
13	The Netherlands	1	39
14	Korea	1	34
	<b>Subtotal</b>	<b>42</b>	<b>2,661</b>
	<b>Grand total</b>	<b>1,546</b>	<b>9,833</b>

# Timelines



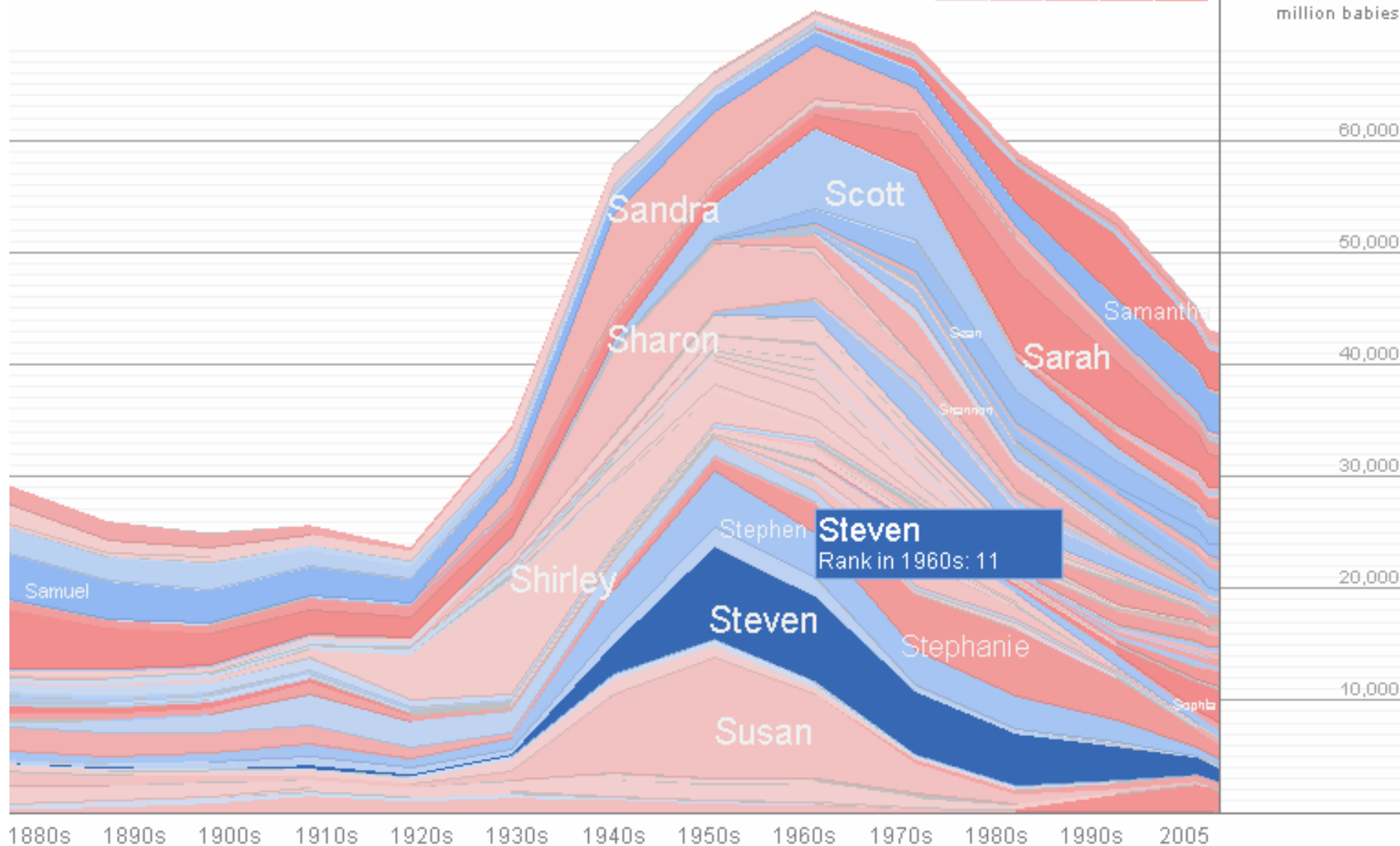
# NameVoyager

>S

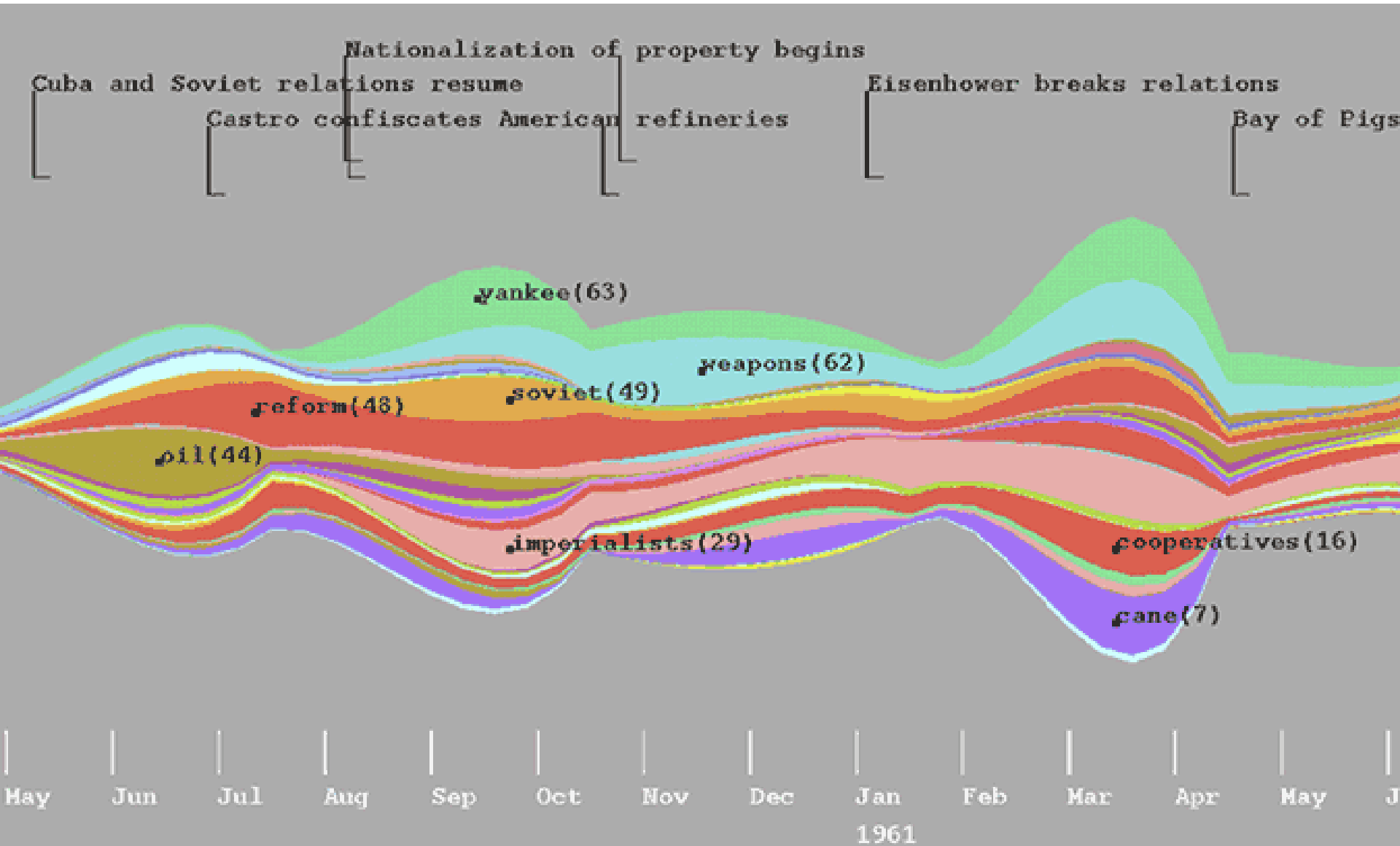
☐ boys
 ☐ girls
 ☒ both

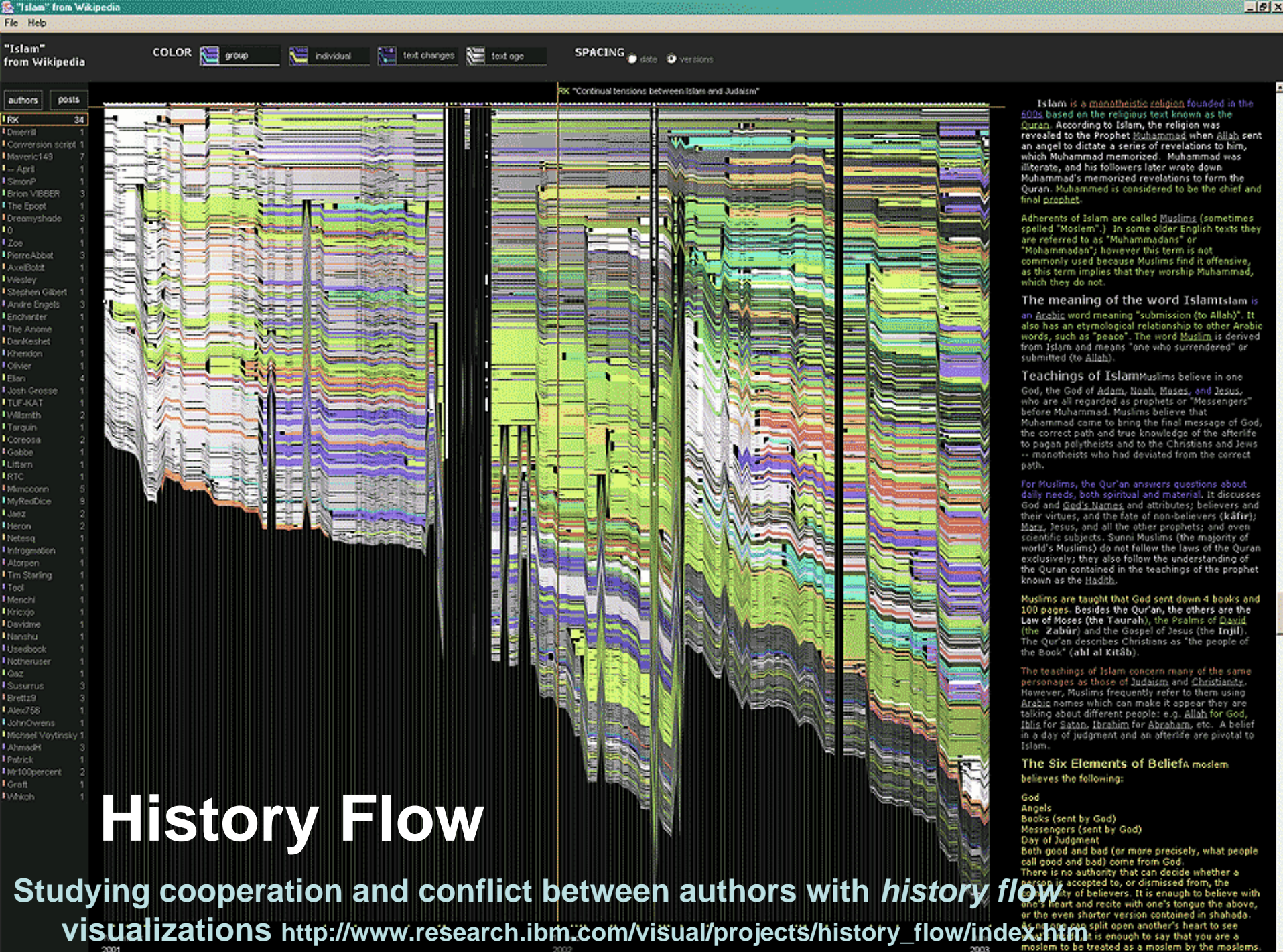
2005 rank, boys	1000	500	100	25	1
girls	1000	500	100	25	1

Names starting with  
"S", per  
million babies



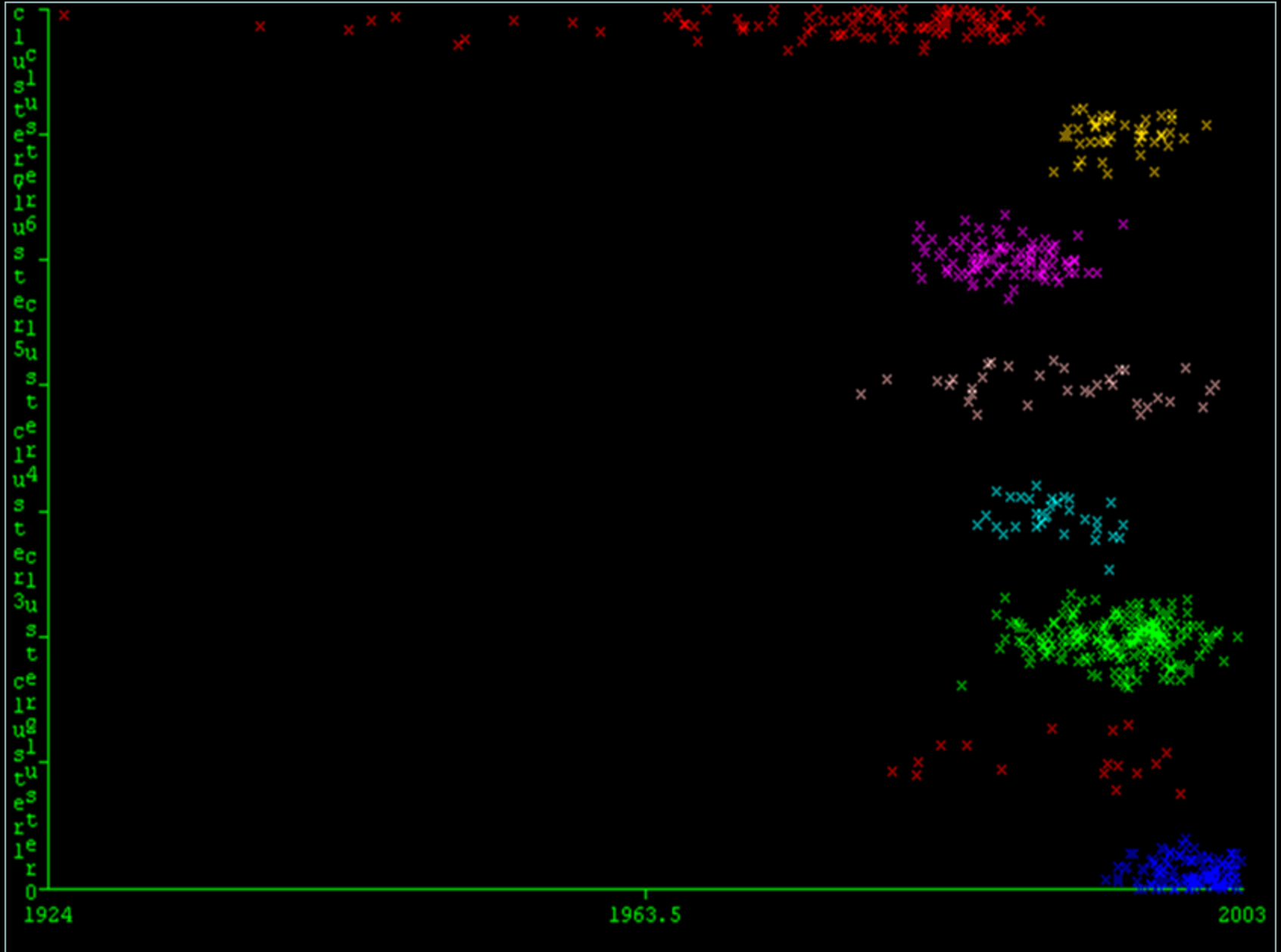
# ThemeRiver







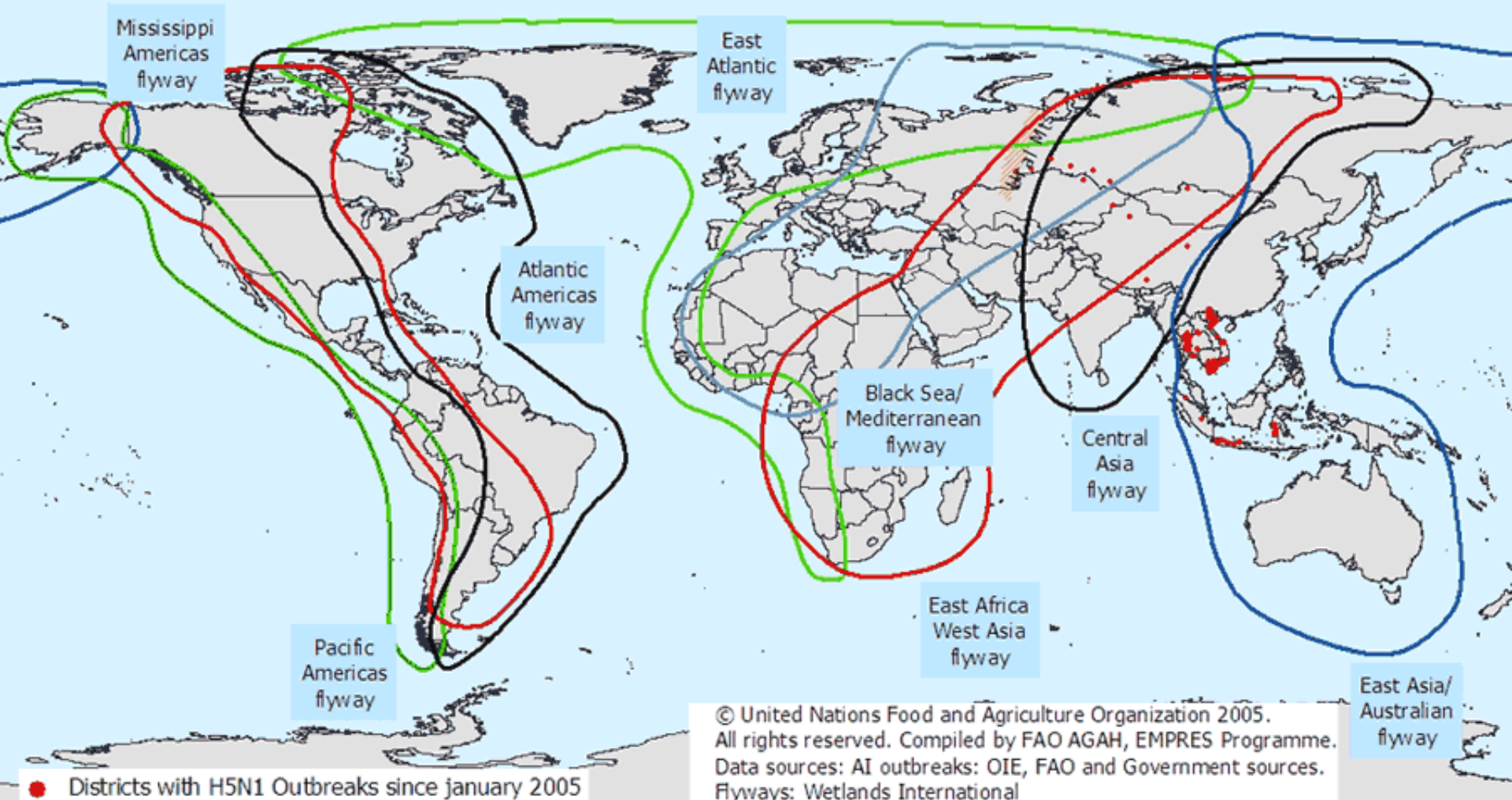
# The Dynamics of Scientific Knowledge



# Spatial and Conceptual Diffusion

## H5N1 outbreaks in 2005 and major flyways of migratory birds

Situation on 30 August 2005





**Geospatial patterns of terrorist events**





Image © 2006 Sanborn  
© 2006 Navteq  
© 2006 TerraMetrics  
© 2006 TerraMetrics  
© 2006 TerraMetrics

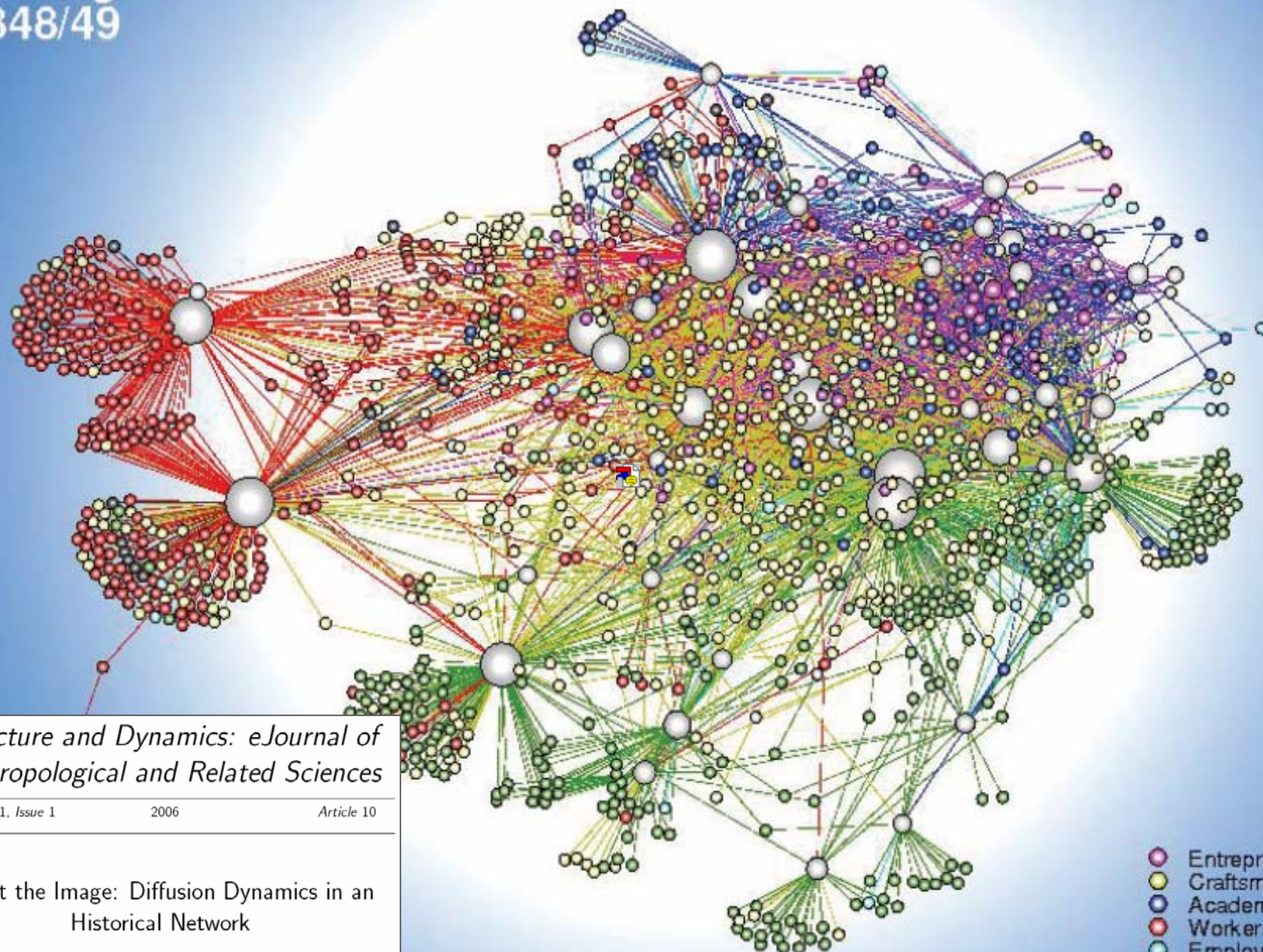
© 2005 Google

# Geospatial patterns of relevant research

Streaming 100%

Eye alt 29537 ft





*Structure and Dynamics: eJournal of  
Anthropological and Related Sciences*

Volume 1, Issue 1

2006

Article 10

About the Image: Diffusion Dynamics in an  
Historical Network

Lothar Krempel\*

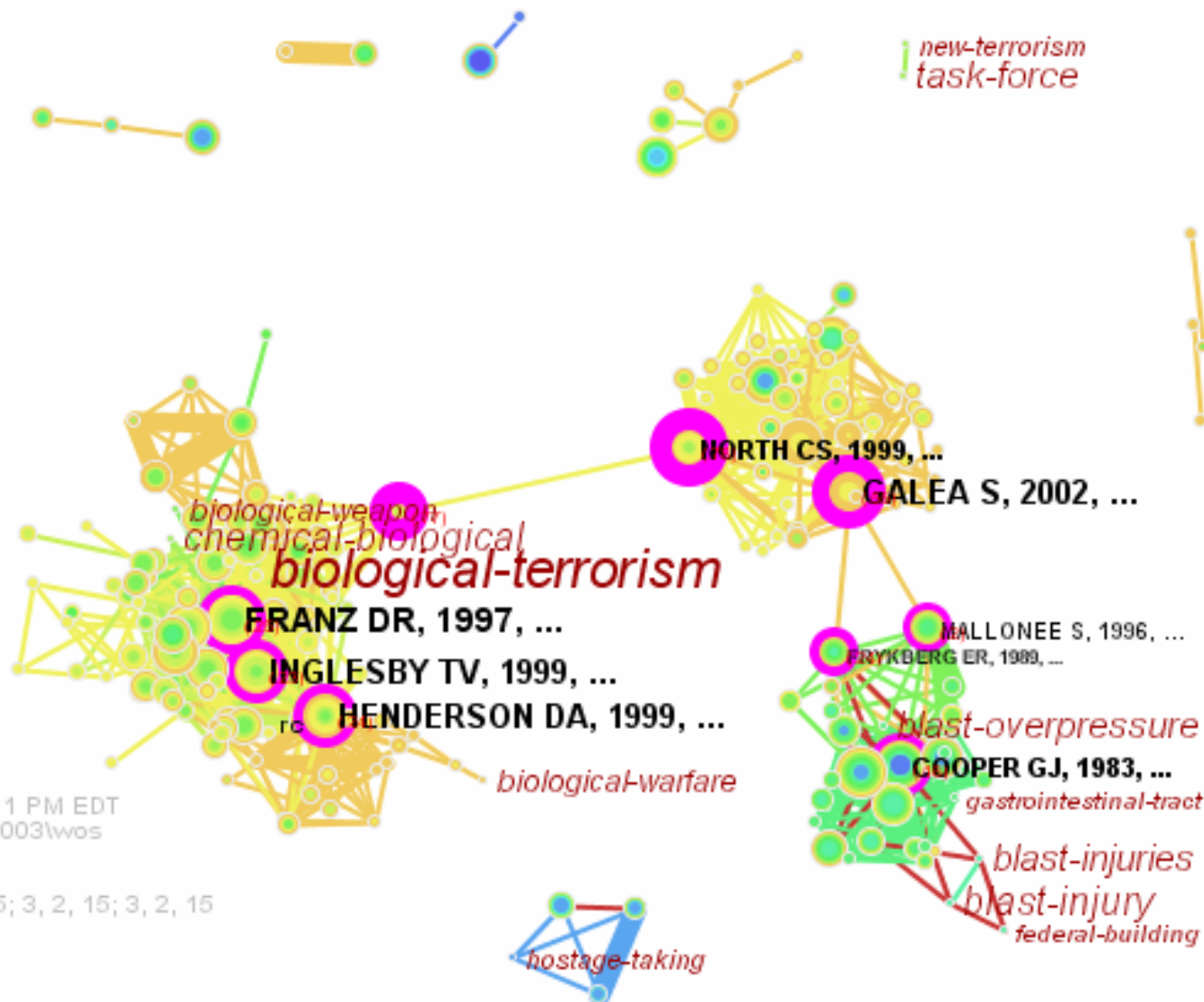
Michael Schnegg†

- Entrepreneurs
- Craftsmen
- Academics
- Workers
- Employees
- Peasants

# The Role of Turning Points in Evolution and ... Diffusion



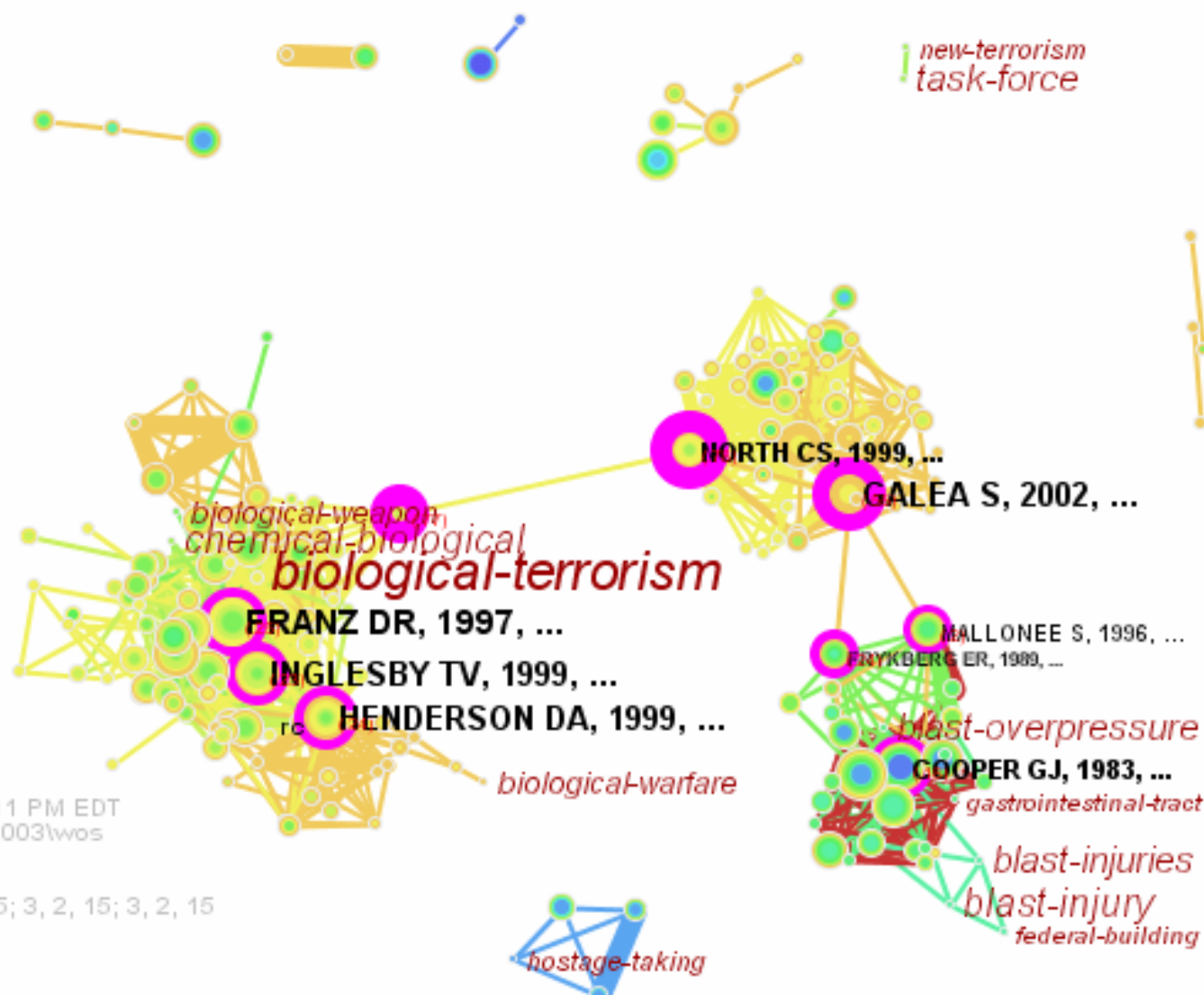
# 1996-1996



CiteSpace, v. 1.0.48  
 Time: October 20, 2005 7:14:11 PM EDT  
 Data: c:\data\terrorism-1990-2003\wos  
 Timespan: 1990-2005  
 Slice Length: 1  
 Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
 Network: N=192, E=873  
 Excluded:

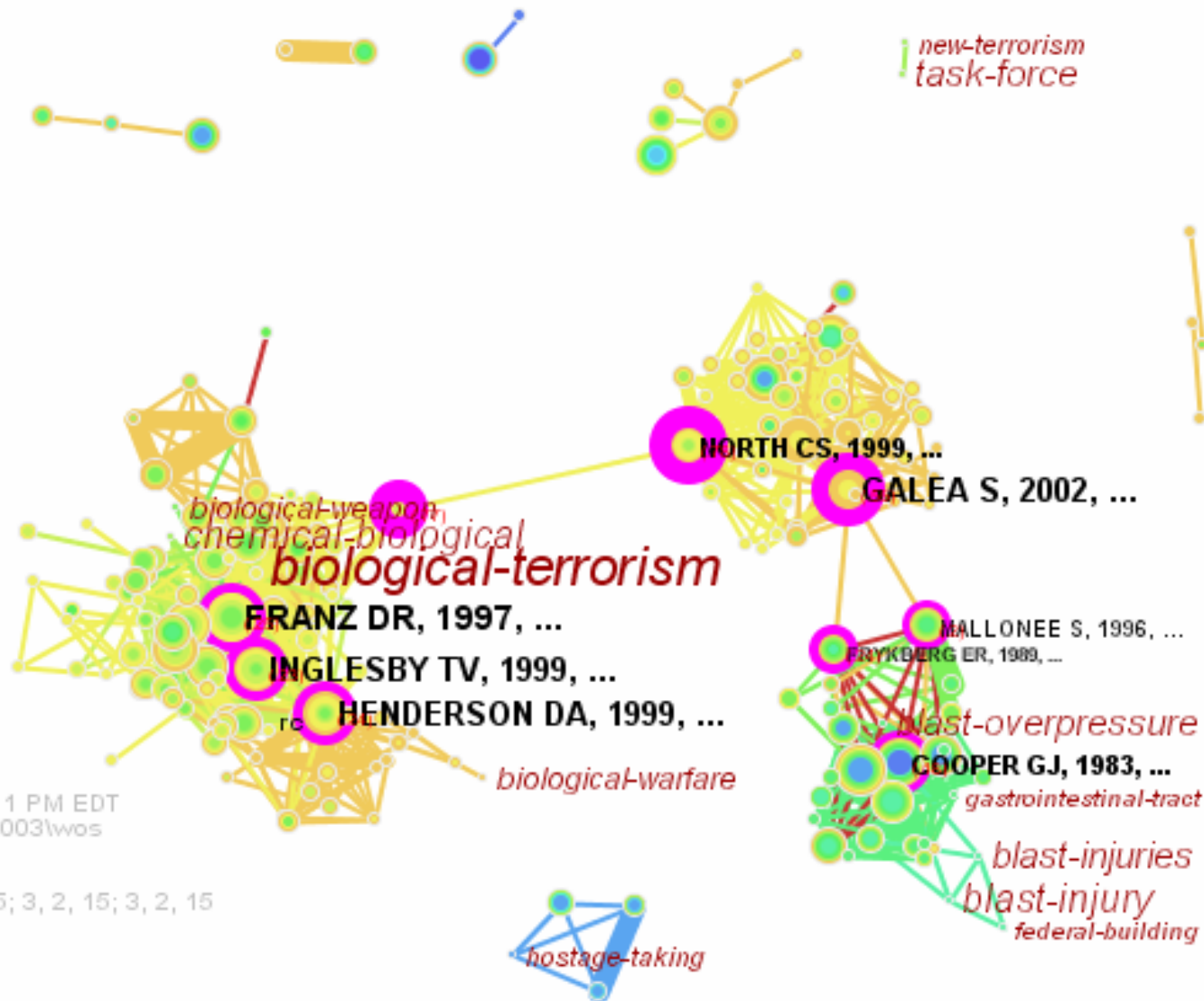


# 1997-1997



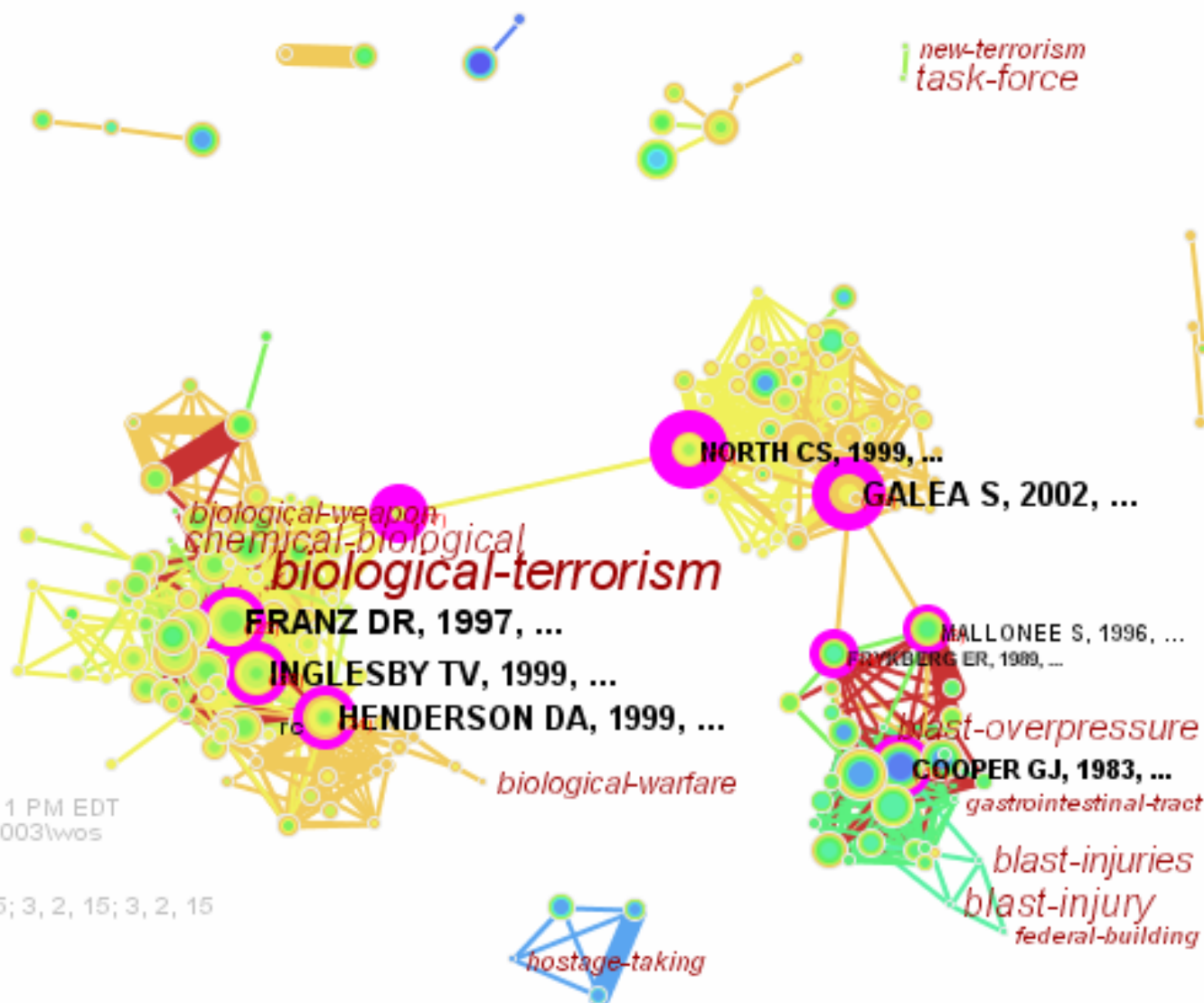
CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

# 1998-1998



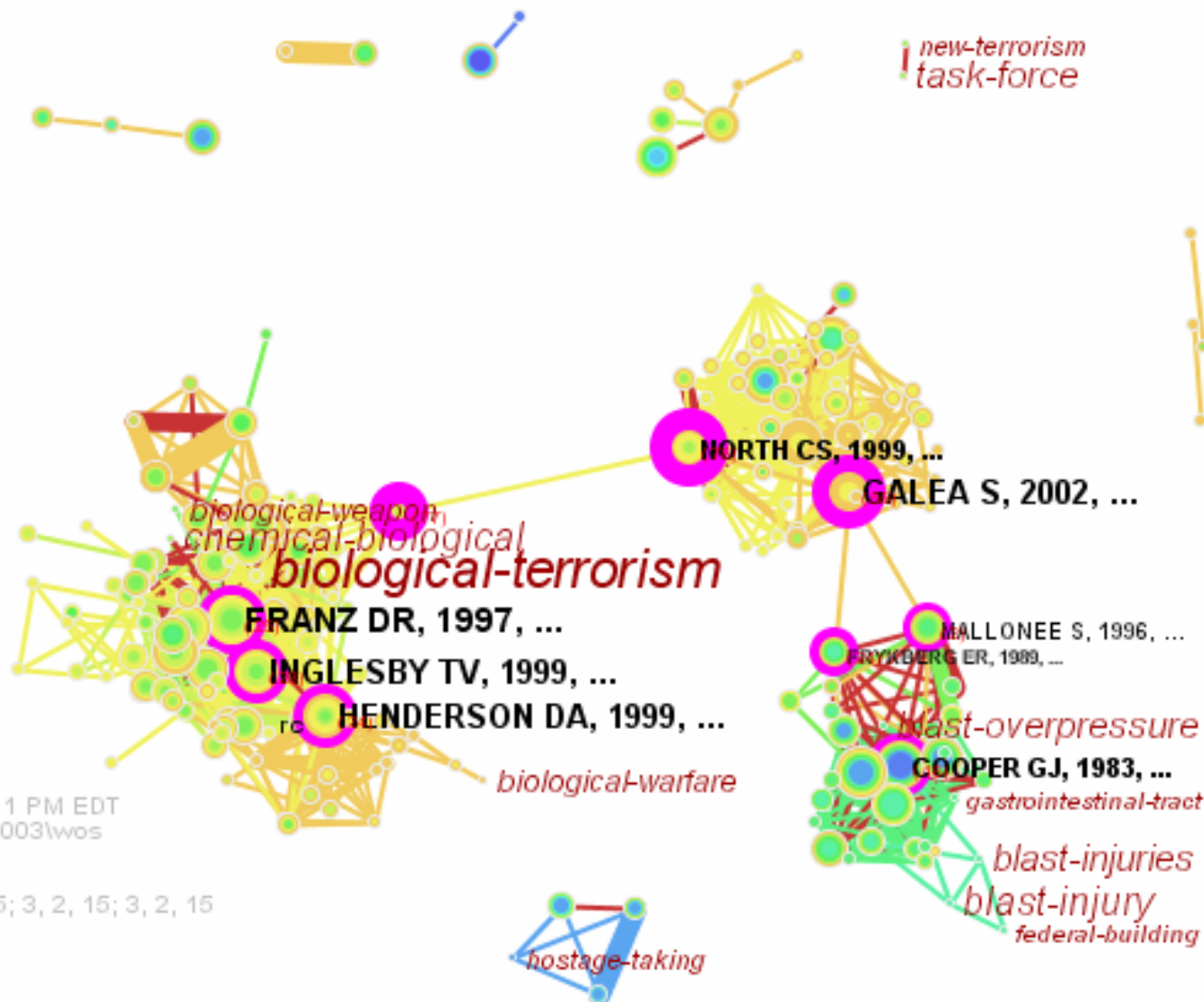
CiteSpace, v. 1.0.48  
 Time: October 20, 2005 7:14:11 PM EDT  
 Data: c:\data\terrorism-1990-2003\wos  
 Timespan: 1990-2005  
 Slice Length: 1  
 Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
 Network: N=192, E=873  
 Excluded:

# 1999-1999



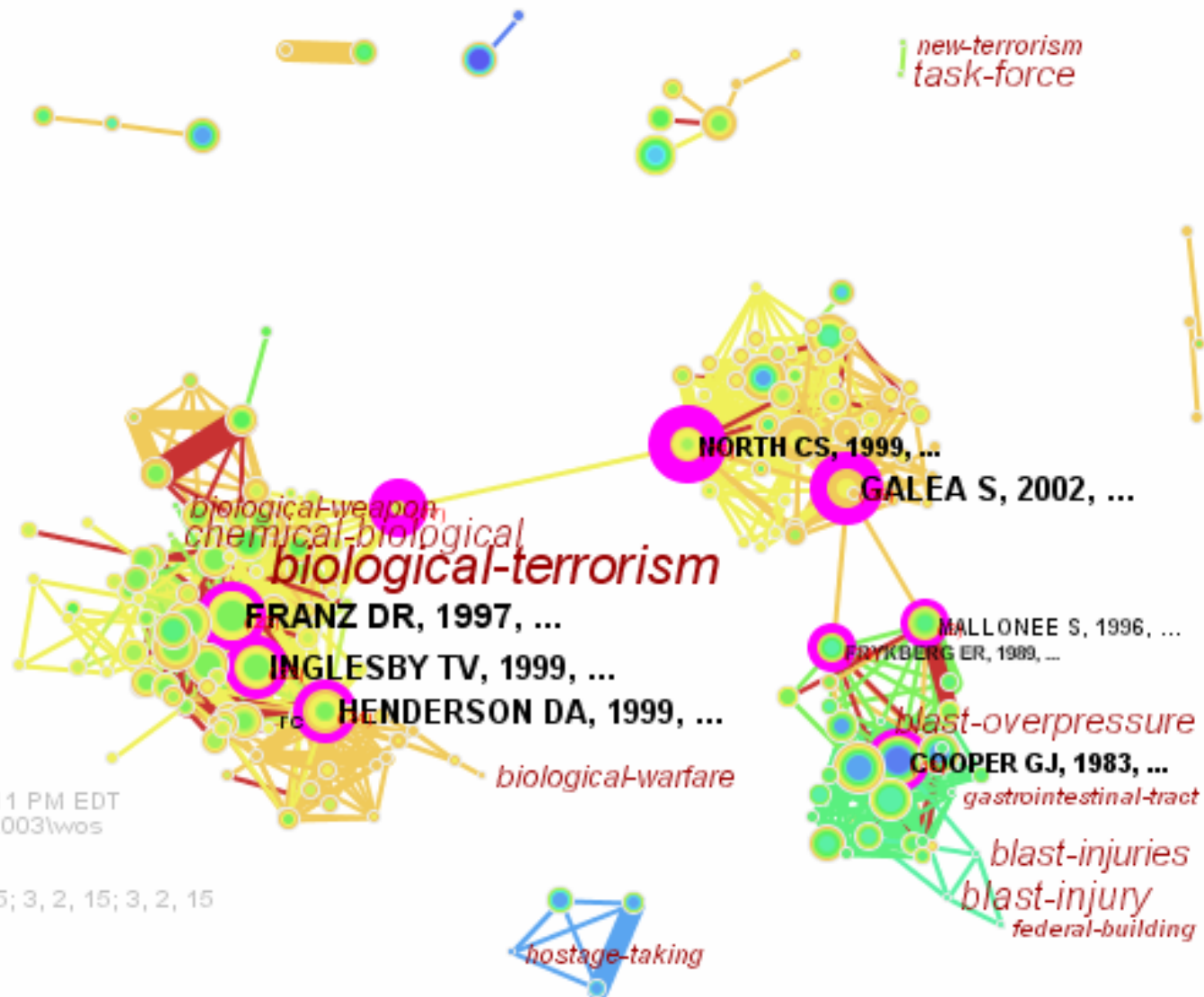
CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

2000-2000



CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

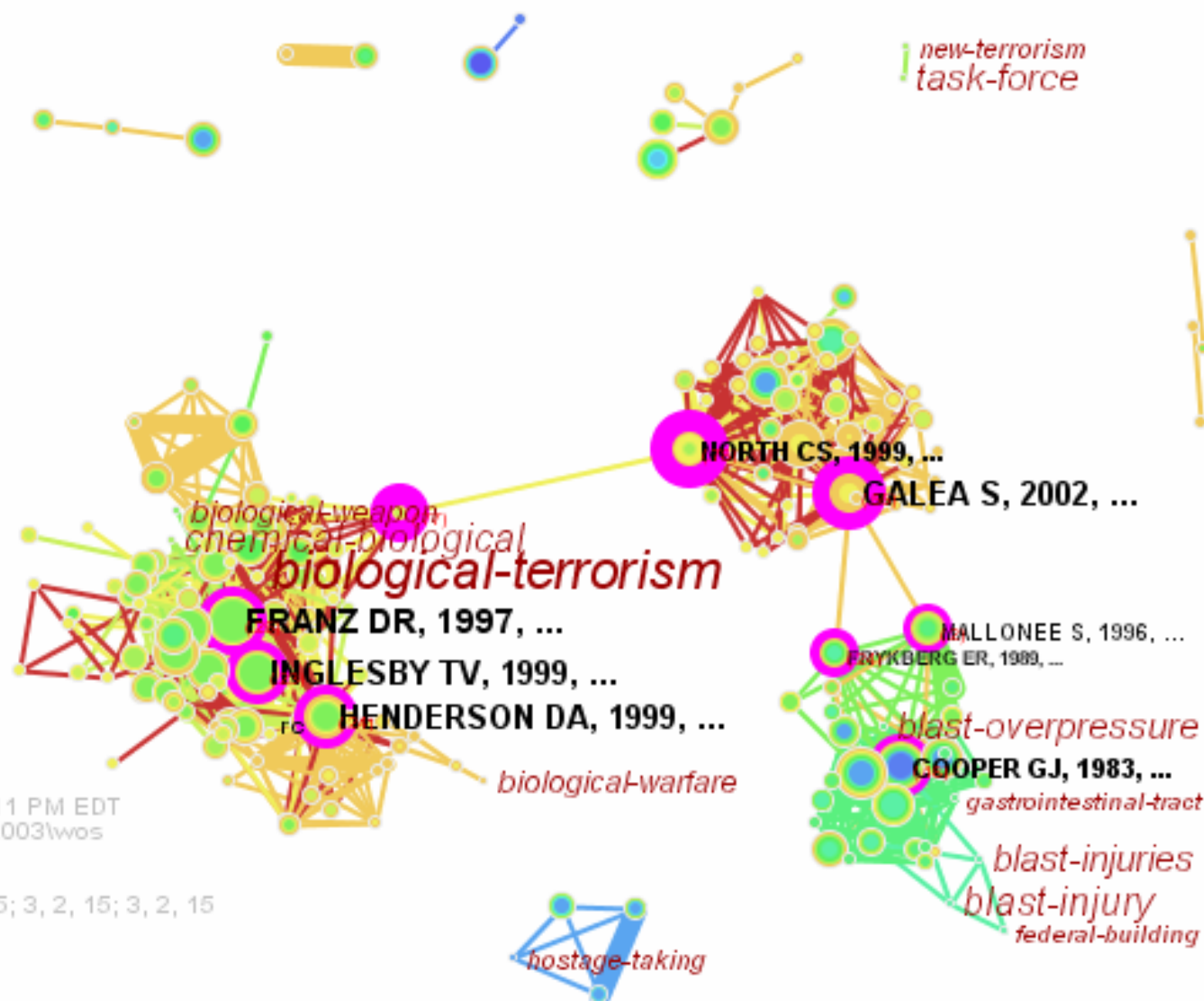
2001-2001



CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

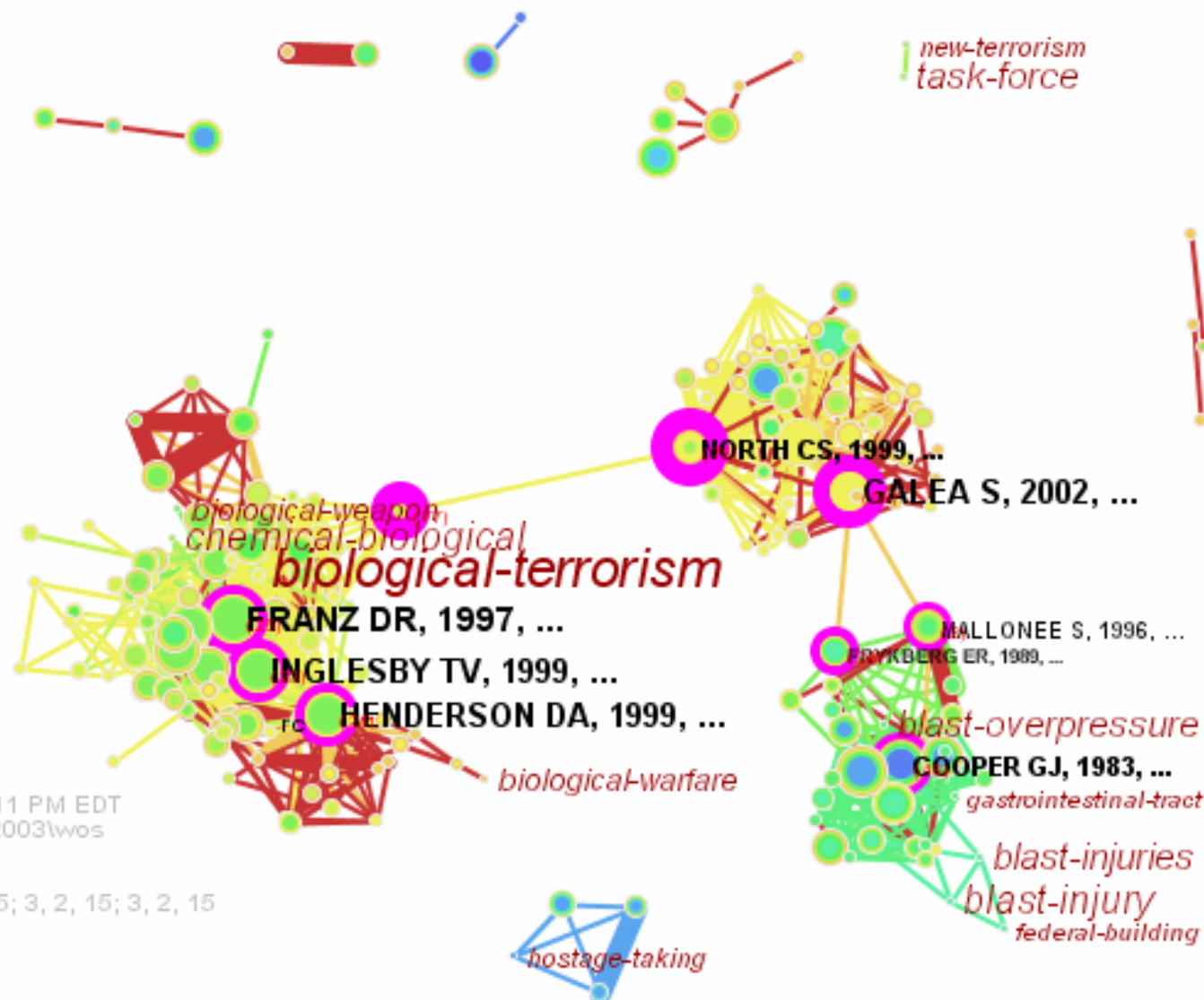


2002-2002



CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

2003-2003



CiteSpace, v. 1.0.48  
Time: October 20, 2005 7:14:11 PM EDT  
Data: c:\data\terrorism-1990-2003\wos  
Timespan: 1990-2005  
Slice Length: 1  
Threshold (c, cc, ccv): 3, 2, 15; 3, 2, 15; 3, 2, 15  
Network: N=192, E=873  
Excluded:

# The Impact of a Pivotal Point

Galea, S. et al. (2002)

How does the citations to this paper  
spread geographically?



2006



USA

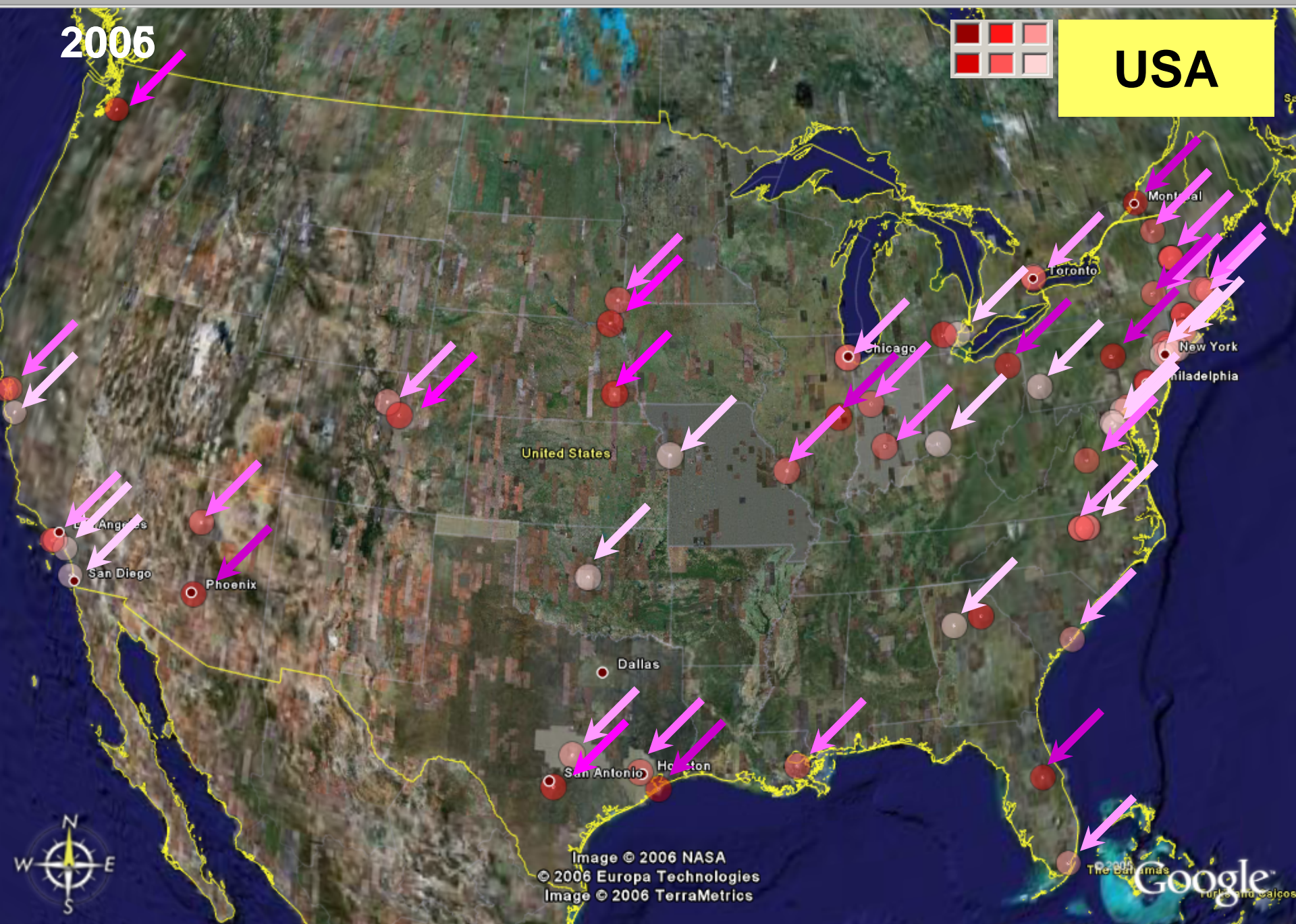


Image © 2006 NASA  
© 2006 Europa Technologies  
Image © 2006 TerraMetrics

Google  
The Bahamas  
Turks and Caicos



# Citers to Galea et al. (2002)





2006



# Europe

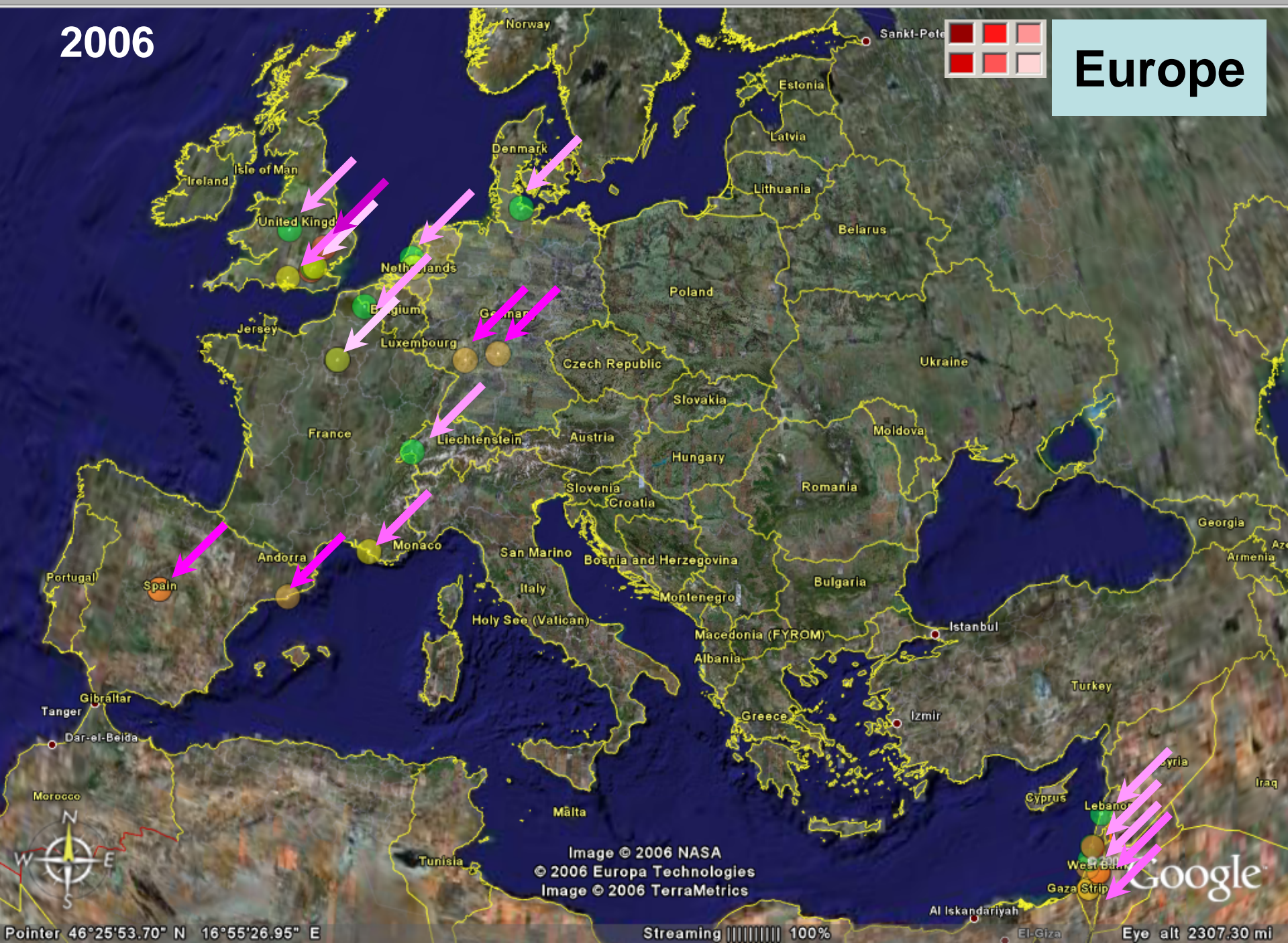


Image © 2006 NASA  
© 2006 Europa Technologies  
Image © 2006 TerraMetrics

Google

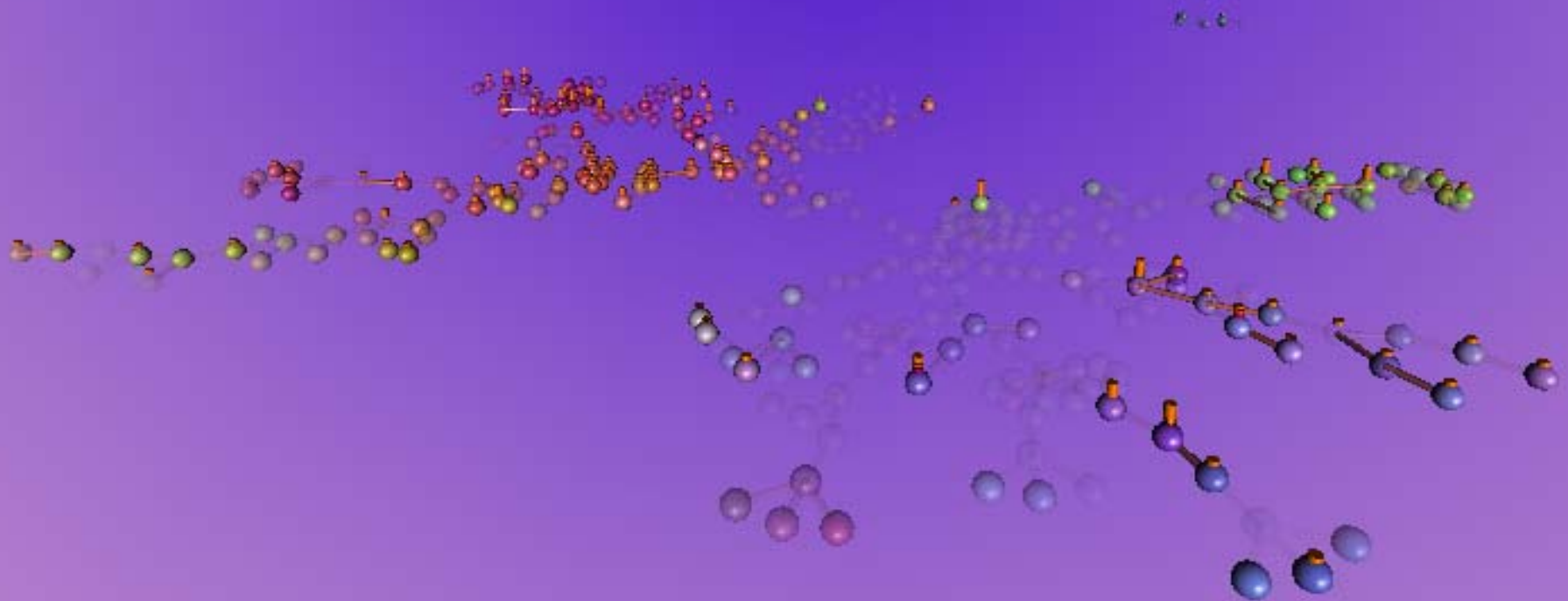
Pointer 46°25'53.70" N 16°55'26.95" E

Streaming 100%

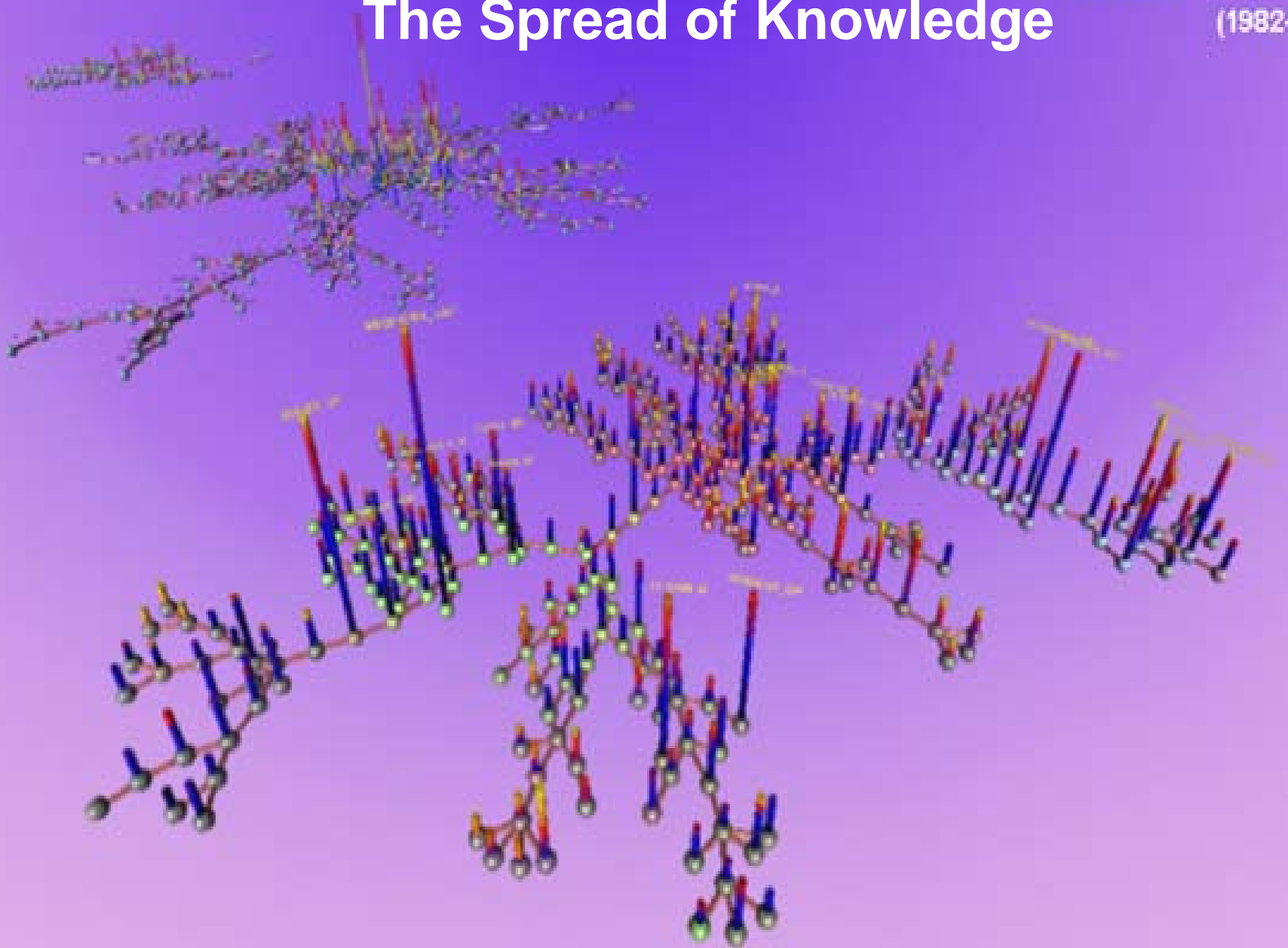
Eye alt 2307.30 mi



# The Spread of Knowledge

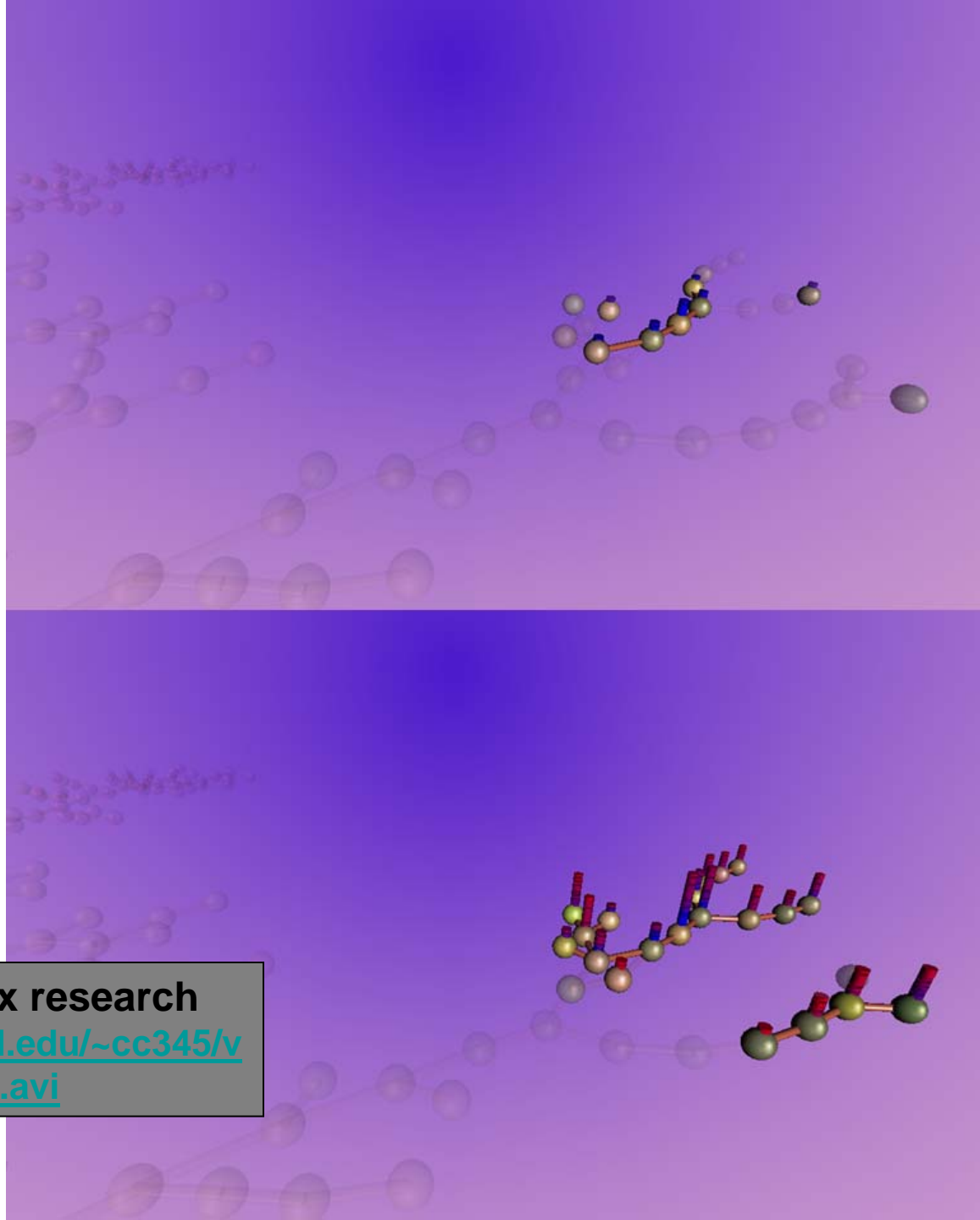


# The Spread of Knowledge



**An animation of botox research**

[http://www.pages.drexel.edu/~cc345/video/citation\\_land\\_local.avi](http://www.pages.drexel.edu/~cc345/video/citation_land_local.avi)



# What influences the diffusion process?

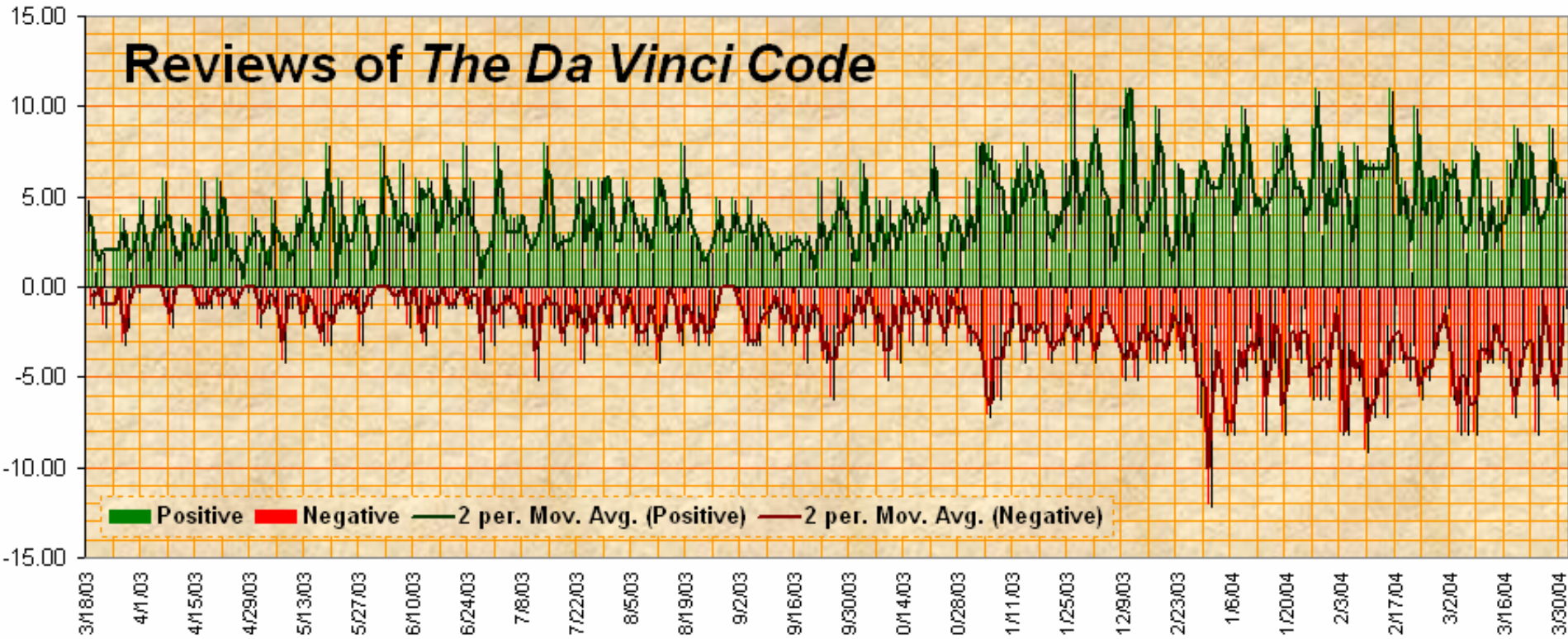
Making Sense of Emergent Social Phenomena

- Amazon
  - Blogmarks
  - CiteULike
  - Del.ici.us
  - ...
  - Wikipedia
- 
- Emotional and Sentimental Analysis
  - Information Diffusion Through Blogspace

# Differentiating Conflicting Views

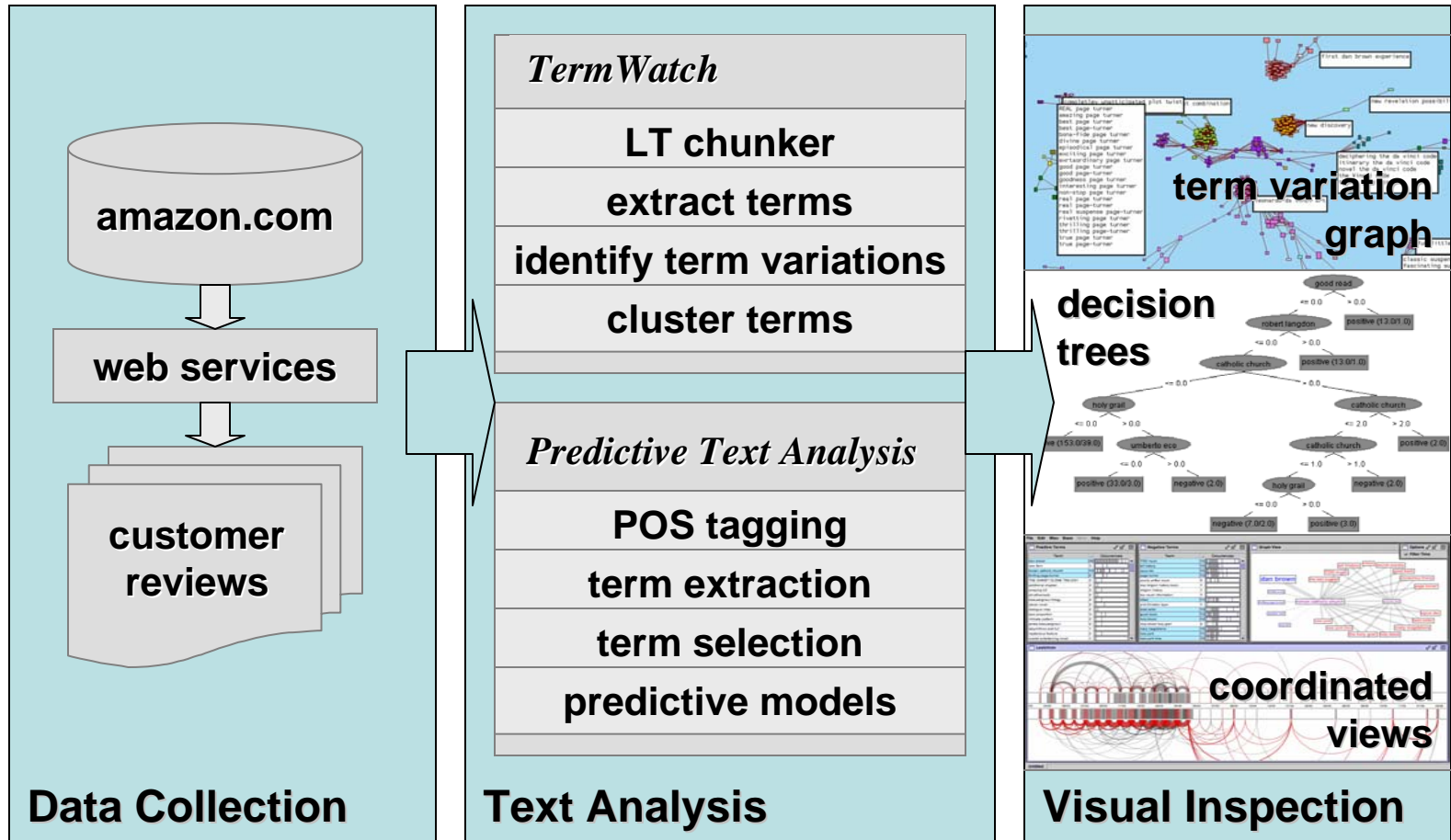
1,738 positive reviews (=4, 5)

918 negative reviews (=1, 2)



[illegible][illegible]





# Term Variations

- Head substitution

age-old religious **secret** → age-old religious **mystery**

- Expansion

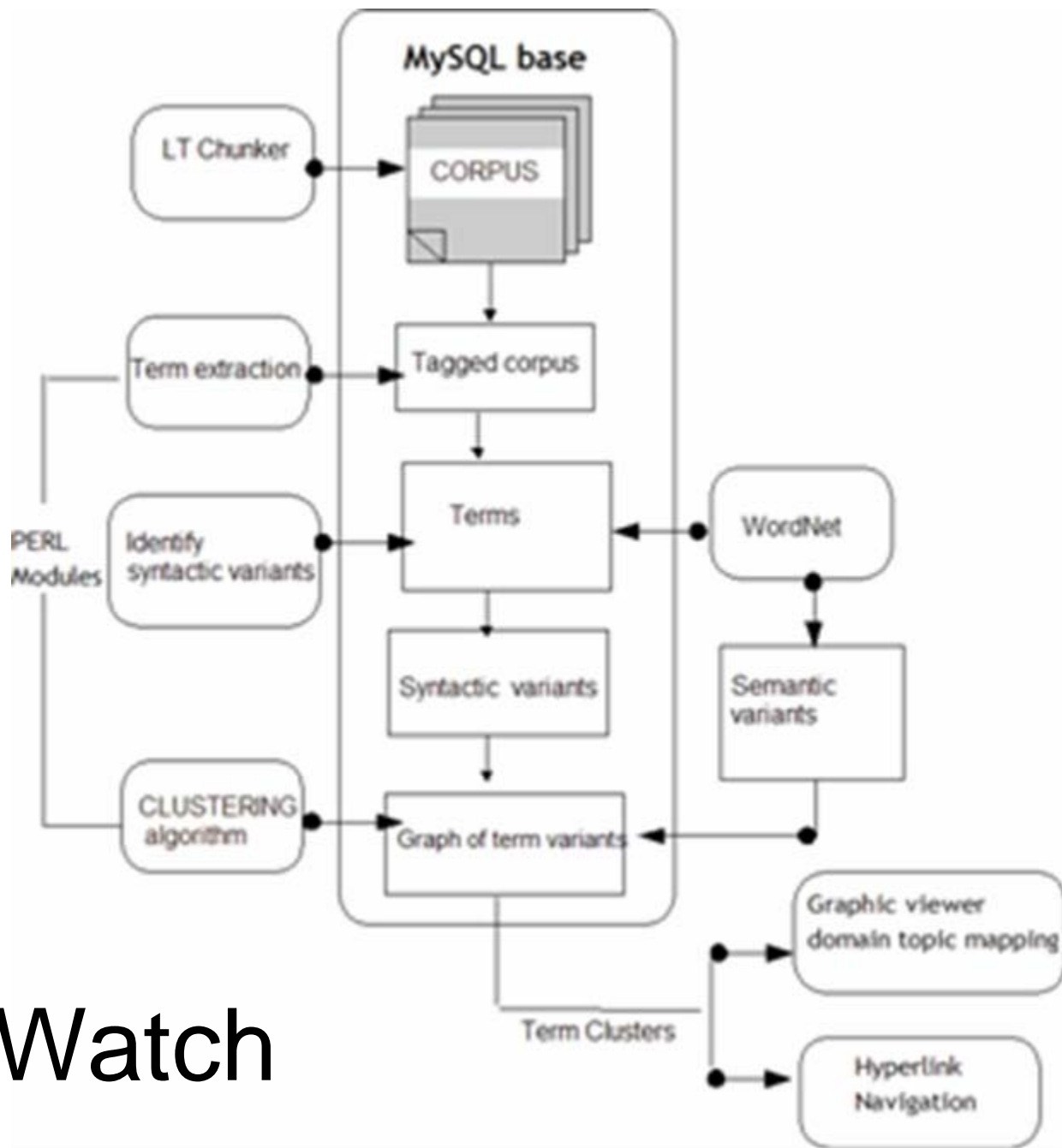
action thriller → **perfect suspense** action thriller

amusing speculation → amusing speculation **collection**

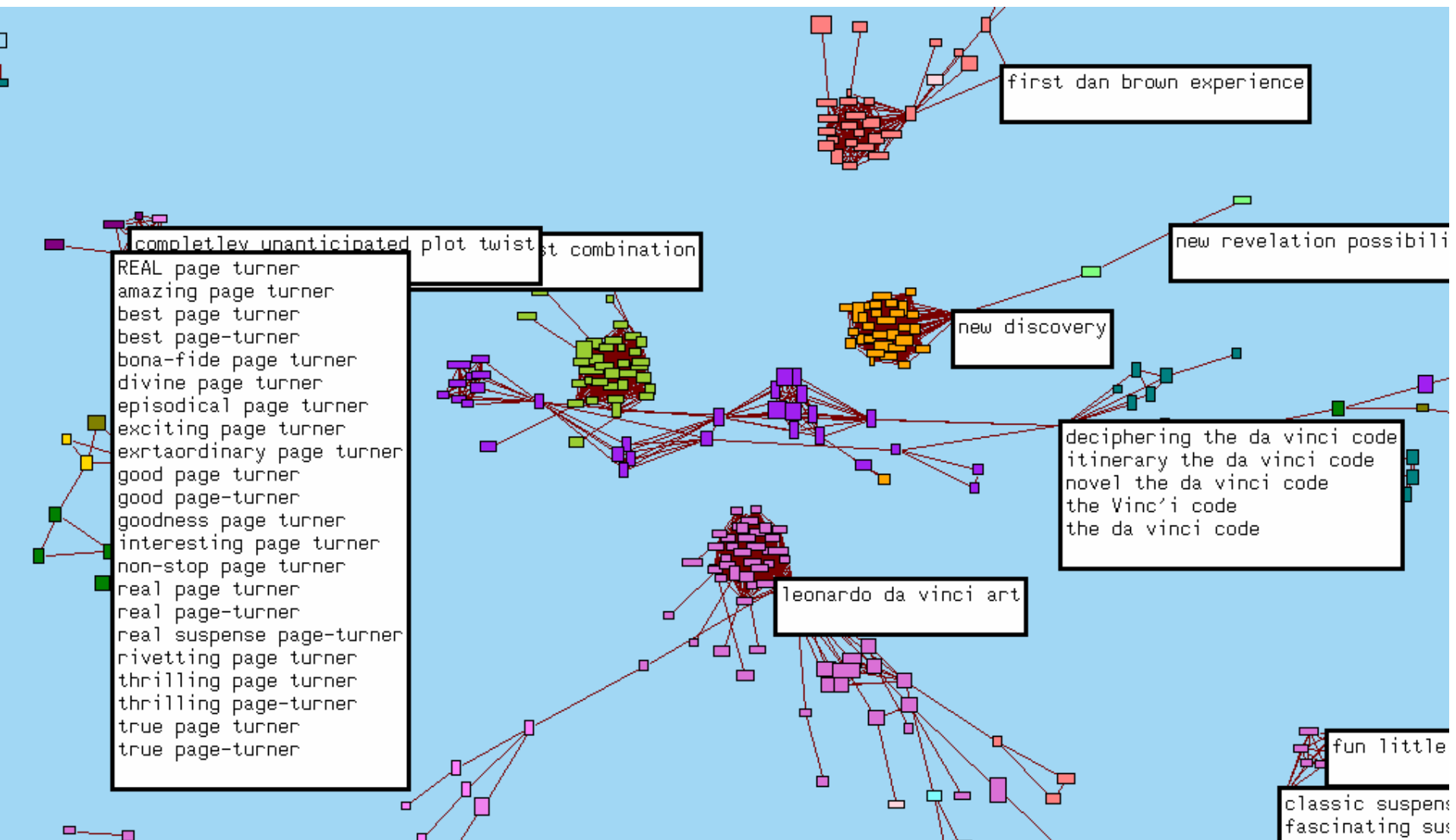
roller coaster → **electrifying** roller coaster **ride**

- Association

mind-binding puzzle ↔ 5-star entertainment

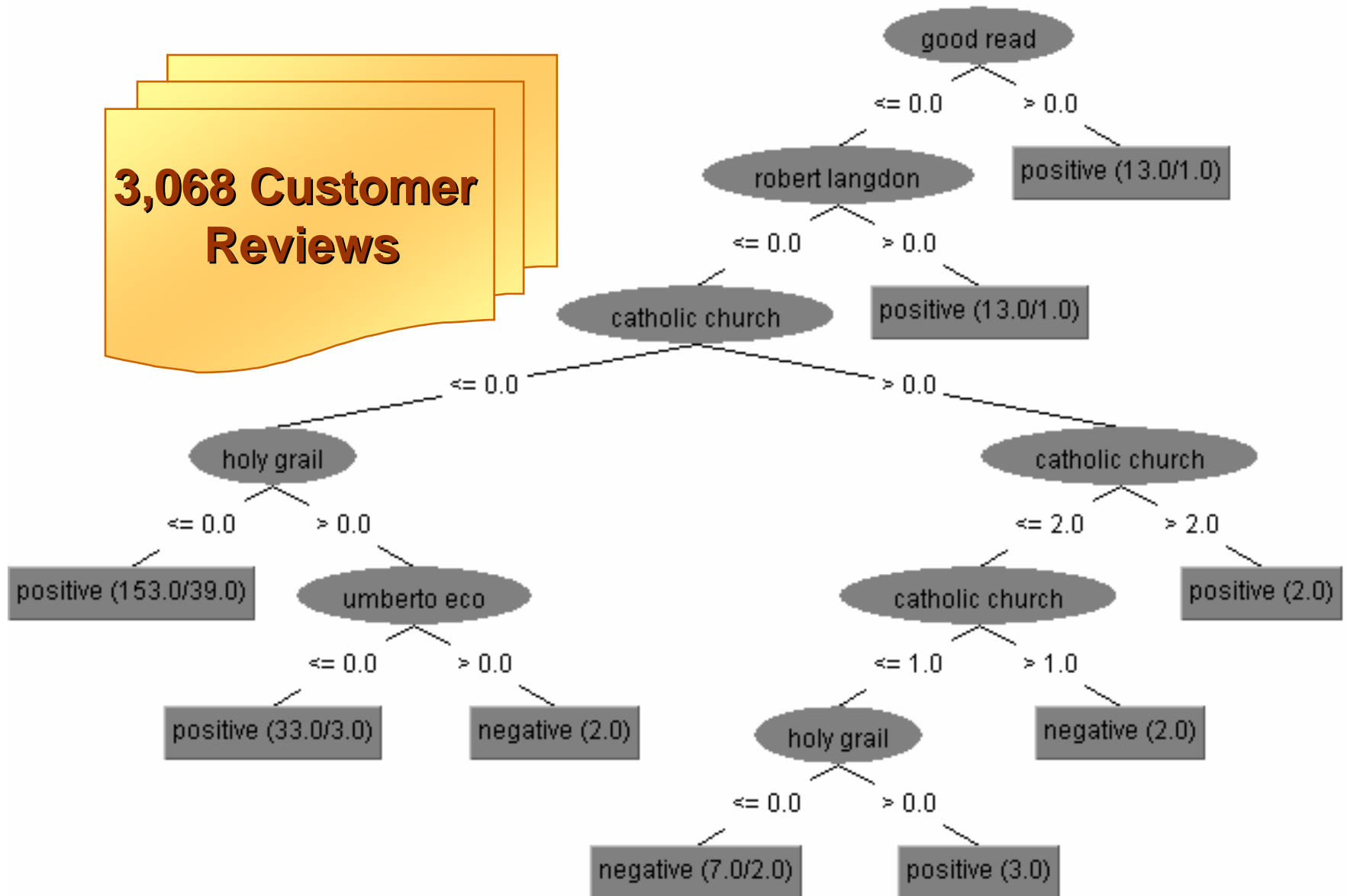


# TermWatch



# *The Da Vinci Code* (Amazon Sales Rank: 7)

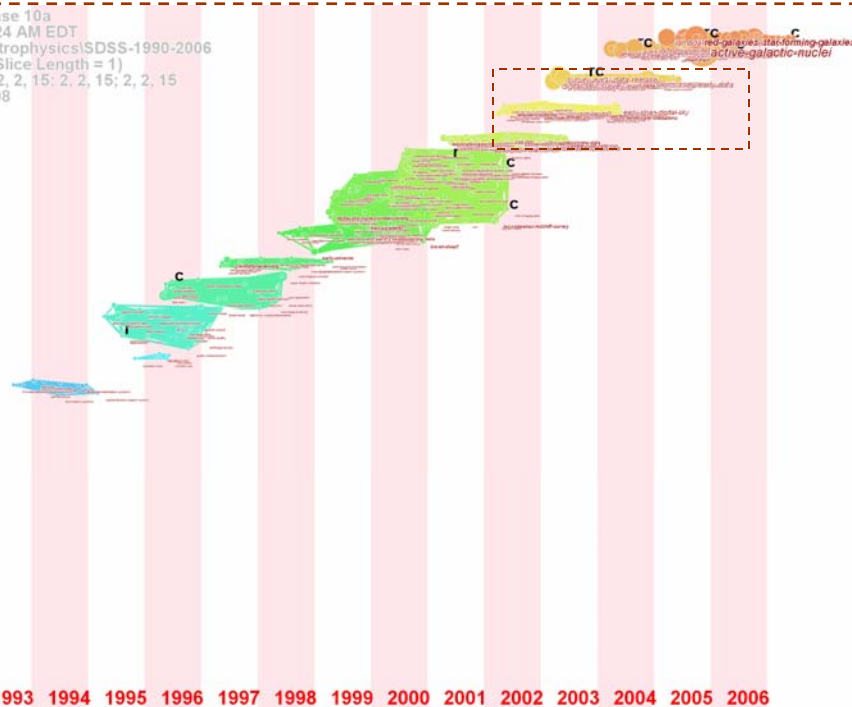
**3,068 Customer Reviews**





# Making Sense of Concept Change

CiteSpace, v. 2.0. Release 10a  
 October 18, 2006 8:06:24 AM EDT  
 C:\DATA\SCIENCE\Astrophysics\SDSS-1990-2006  
 Timespan: 1990-2006 (Slice Length = 1)  
 Threshold (c, cc, ccv): 2, 2, 15; 2, 2, 15; 2, 2, 15  
 Network: N=404, E=4508



survey-early-data-release  
 digital-sky-survey-early-data  
 similar-to  
 Sloan-Digital-Sky-Survey  
 smaller-mass-functions

angular-clustering  
 spectrum  
 similar-to  
 early-sloan-digital-sky  
 monte-carlo-simulations  
 sdss-1  
 magnitude-bin  
 angular-clustering  
 spectrum  
 similar-to  
 early-sloan-digital-sky  
 monte-carlo-simulations  
 sdss-1

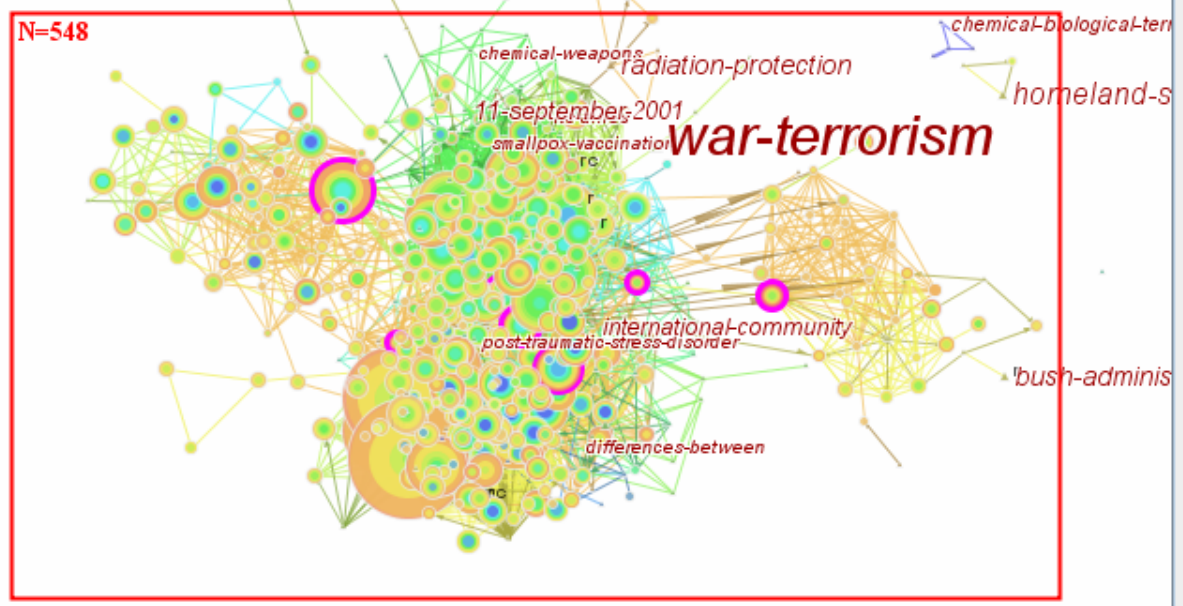
100-000  
 absorption-line  
 line-broad-absorption-line  
 commissioning-data  
 sdss-1  
 complete  
 data

2001-2003



Step Up [Navigation icons] Text search ...

CiteSpace, v. 2.0. Release 10b  
October 22, 2006 5:33:29 PM EDT  
C:\DATA\Terroris- 1990-2006.3.21  
Timespan: 1996-2006 (Slice Length = 1)  
Threshold (c, cc, cov): 3, 2, 15; 4, 4, 20; 5, 4, 20  
Network: N=565, E=4673



Search Results Quick Guide

Control Panel

Links to PubMed

- [12226151] PREZANT DJ, 2002
- [15043650] GIGERENZER G, 2004
- [3892222] EVANS ME, 1985
- [7013615] BRACHMAN PS, 1980
- [9696731] SHAPIRO RL, 1998
- [11923491] BREMAN JG, 2002
- [11925351] BARTHOLOMEW RE, 2002
- [2003249] KLEIN JS, 1991
- [8460135] ABRAMOVA FA, 1993
- [10559102] SHAFAZAND S, 1999
- [10591317] FRIEDLANDER AM, 1999
- [11740711] GLASS TA, 2002
- [15197022] WASELENKO JK, 2004
- [15740448] MOGHADDAM FM, 2005
- [11704685] BUSH LM, 2001
- [8486963] FRIEDLANDER AM, 1993
- [8414405] MADER TH, 1993
- [8689247] QUENEMOEN LE, 1996
- [9132196] CARLEY SD, 1997
- [11030310] KHAN AS, 2000
- [11316014] LOEWENSTEIN GF, 2001
- [3563507] SLOVIC P, 1987

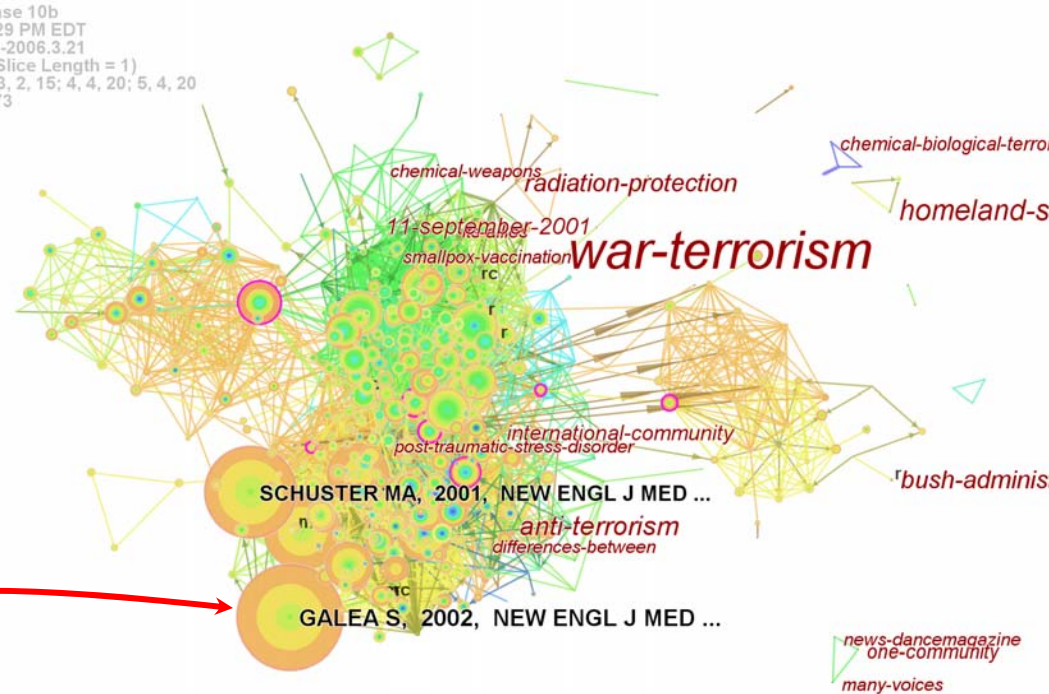
Node Details		MeSH Headings - Major Topics		MeSH Headings - Minor Topics						
Freq	Burst	Centrality	Keyword	Author	Year	Title	Source	Vol	Page	HalfLife
149		0.03		GALEA S	2002	Psychological	NEW ENGL J MED	V346	P982	0
128		0.01		SCHUSTER MA	2001	A national survey of	NEW ENGL J MED	V345	P1507	1
106		0.02		SCHLENGER WE	2002	Psychological	JAMA-J AM MED ASSOC	V288	P581	1
85		0.02		SILVER RC	2002	Nationwide	JAMA-J AM MED ASSOC	V288	P1235	1
77		0.01	war-terrorism							
70		0.04		*AM PSYCH ASS	1994	Title not found	DIAGN STAT MAN ME...	VBOOK	P0	3
62		0.09		NORTH CS	1999	Psychiatric disorders	JAMA-J AM MED ASSOC	V282	P755	3

## GALEA 2002

WoS: **149** Citations (as of 3/21/2006 within the terrorism data set)

WoS: **292** Citations (as of 10/22/2006)

Google Scholar: **308** Citations (as of 10/22/2006)



CiteSpace - Search Results

Google Scholar BETA

Search "GALEA S" Advanced Scholar Search Scholar Preferences Scholar Help

Scholar All articles Recent articles Results 1 - 10 of about 1,270 for "GALEA S". (0.12 seconds)

All Results

[S Galea](#)  
[D Vlahov](#)  
[J Ahern](#)  
[H Resnick](#)  
[D Kilpatrick](#)

[CITATION] Psychological Sequelae of the September 11 Terrorist Attacks in New York City - [group of 5 »](#)  
 S Galea, J Ahern, H Resnick, D Kilpatrick, M ...  
 Background The scope of the terrorist attacks of September 11, 2001, was unprecedented in the United States. We assessed the prevalence and ...  
 Cited by 308 - [Related Articles](#) - [Web Search](#) - [BL Direct](#)

[Trends of Probable Post-Traumatic Stress Disorder in New York City after the September 11 Terrorist ... - group of 6 »](#)  
 S Galea... - American Journal of Epidemiology, 2003 - sph.umich.edu  
 Page 1. 514 Am J Epidemiol 2003;158:514-524 American Journal of Epidemiology  
 Copyright © 2003 by the Johns Hopkins Bloomberg School ...  
 Cited by 74 - [Related Articles](#) - [View as HTML](#) - [Web Search](#) - [BL Direct](#)



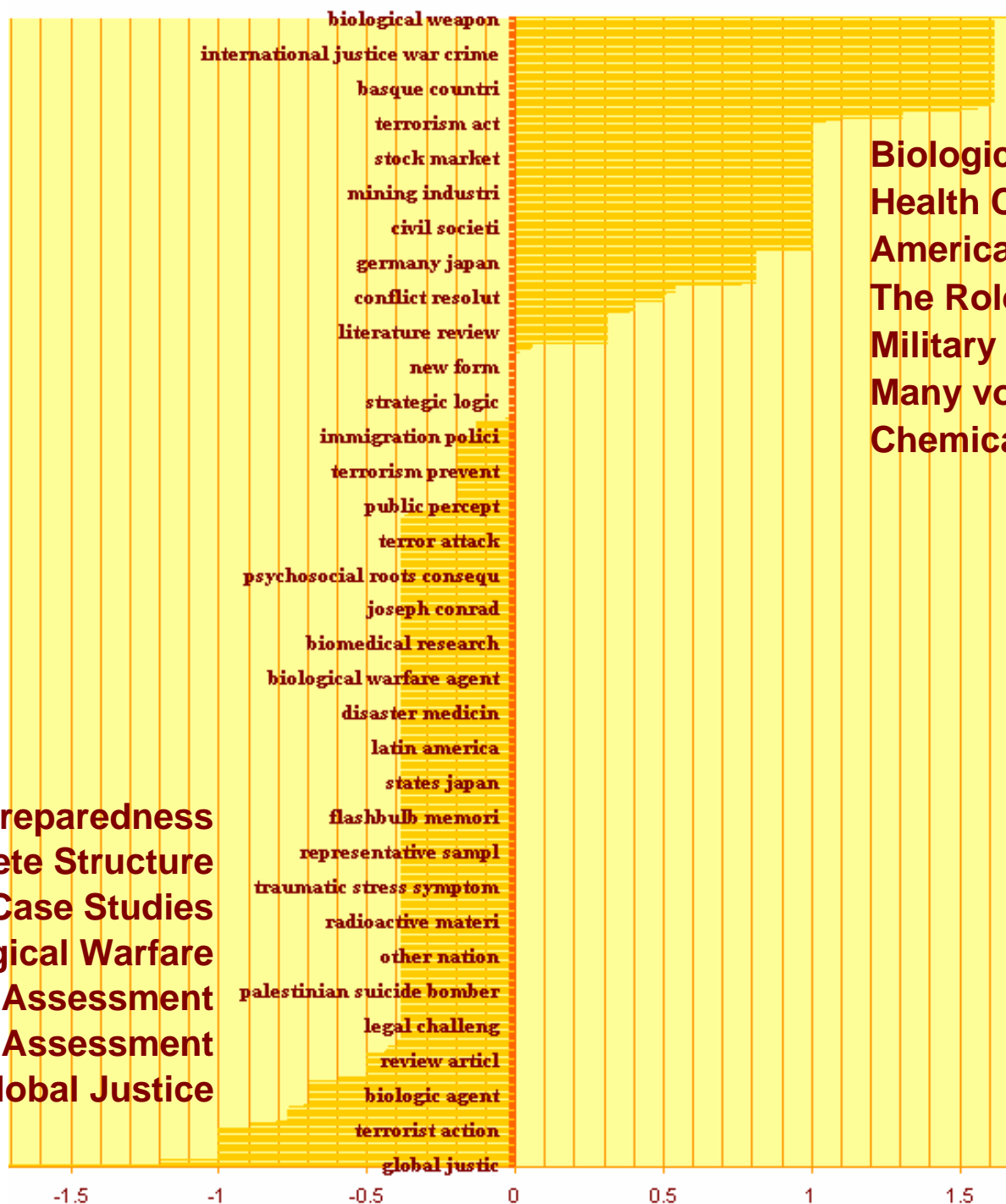




2002	2003	2004	2005	2006
25 new york city	38 posttraumatic stress disorder	39 posttraumatic stress disorder	32 world trade center	40 posttraumatic stress disorder
22 posttraumatic stress disorder	27 new york city	25 world trade center	26 traumatic stress disorder	38 world trade center
11 risk factor	20 world trade center	14 mental health service	21 new york city	33 new york city
10 world trade center	14 va patient	14 new york city	13 posttraumatic stress disorder	14 project liberty counseling
10 health effect	13 world trade	14 world trade	13 probable posttraumatic stress disorder	14 alcohol use
9 ptsd symptom	12 odds ratio	14 traumatic stress disorder	10 media exposure	14 traumatic stress disorder
9 world trade	12 risk factor	12 risk factor	10 representative sample	12 project liberty
9 national comorbidity survey	9 veterans affair	10 mental health problem	9 drug use	12 project liberty service
7 drug user	9 telephone survey	10 fine particulate matter	9 ptsd symptom	11 mental health service
7 disaster site	9 health care	10 mental health impact	8 mental health service	11 world trade
7 vietnam veteran	9 ptsd symptom	9 acute stress disorder	8 new york city resident	11 family member
7 adult resident	8 vietnam veteran	8 language use	6 risk factor	10 project liberty crisis
6 confidence interval	8 substance use	8 medication use	6 alcohol consumption	9 community disaster
6 odds ratio	8 national comorbidity survey	7 health service use		9 negative life event
6 main outcome measure	8 panic attack	6 mental health visit		8 service recipient
	8 acute stress disorder	6 health care system		8 alcohol dependence
	8 vehicle accident	6 vehicle accident		8 binge drinking
	7 traumatic stress disorder	6 college student		8 wave panel study
	6 symptom severity	6 minority statu		8 mental health statu
		6 media coverage		8 telephone interview
		6 ptsd symptom		7 new york city adult
				7 project liberty client
				7 risk factor
				7 center disaster
				6 odds ratio
				6 health service
				6 indirect trauma exposure
				6 ptsd symptom

2002-2003

Emergency Preparedness  
Concrete Structure  
Case Studies  
Biological Warfare  
Risk Assessment  
Survey Assessment  
Global Justice



Biological Weapon  
Health Care  
America's War  
The Role  
Military Strategy  
Many voice  
Chemical Weapon

2005-2006

# Concluding Remarks

- Understanding the dynamics of scientific knowledge and emergent societal information is an increasingly common need across a wide range of domains.
- We need to understand the process itself better in order to build more effective tools.
- We need to take a holistic approach to address these issues from societal as well as individual perspectives.
- We need a critical and sustained mass to achieve this ambitious and challenging mission.
- This is a potential broad-impact area for making real changes to the way we work with large amounts of information.

# Acknowledgements



- **NSF IIS Award #0612129**
  - **SEI: Coordinated Visualization and Analysis of Sky Survey Data and Astronomical Literature**
- **National Visualization and Analytics Center (NVAC)**
  - **Northeast Visualization and Analytics Center (NEVAC)**
  - **Knowledge-enabled Visual Analytics: Supporting Individuals and Terms from Analysis through Action**



# Background Readings

Handout of this Presentation:

- Chen, C. (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377.
- Chen, C. (2004) Searching for intellectual turning points: Progressive Knowledge Domain Visualization. *Proceedings of the National Academy of Sciences of the United States of America* (PNAS), 101 (Suppl. 1), 5303-5310
- Chen, C. (2004) *Information Visualization: Beyond the Horizon*. Springer. 2nd ed. ISBN: 1-85233-789-3.
- Chen, C. (2003) *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*. Springer. ISBN: 1-85233-494-0.

